Optimality conditions and complementarity, Nash equilibria and games, engineering and economic application

Michael C. Ferris

University of Wisconsin, Madison

Funded by DOE-MACS Grant with Argonne National Laboratory

NATCOR, University of Edinburgh, June 16, 2016
Convex subdifferentials

Assume $f$ is convex, then
$$f(z) \geq f(x) + \nabla f(x)^T (z - x)$$
(linearization is below the function)

Incorporate constraints by allowing $f$ to take on $+\infty$ if constraint is violated

$$f : \mathbb{R}^n \mapsto (-\infty, +\infty]$$

$$\partial f(x) = \{ g : f(z) \geq f(x) + g^T (z - x), \forall z \},$$
the subdifferential of $f$ at $x$
Convex subdifferentials

Assume $f$ is convex, then
\[ f(z) \geq f(x) + \nabla f(x)^T (z - x) \]
(linearization is below the function)

Incorporate constraints by allowing $f$ to take on $+\infty$ if constraint is violated

If $f$ is differentiable and convex, then $\partial f(x) = \{ \nabla f(x) \}$
e.g. $f(z) = \frac{1}{2} z^T Q z + p^T z$, then $\partial f(x) = \{ Q x + p \}$
$x^*$ solves $\min f(x)$ if and only if $0 \in \partial f(x^*)$
Indicator functions and normal cones

\[ \psi_C(z) = \begin{cases} 0 & \text{if } z \in C \\ \infty & \text{else} \end{cases} \]

\( \psi_C \) is a convex function when \( C \) is a convex set
Indicator functions and normal cones

\[ \psi_C(z) = \begin{cases} 
0 & \text{if } z \in C \\
\infty & \text{else}
\end{cases} \]

\( \psi_C \) is a convex function when \( C \) is a convex set

If \( x \in C \), then

\[ g \in \partial \psi_C(x) \iff \psi_C(z) \geq \psi_C(x) + g^T(z - x), \quad \forall z \]

\[ 0 \geq g^T(z - x), \quad \forall z \in C \]

Normal cone to \( C \) at \( x \),

\[ N_C(x) := \partial \psi_C(x) = \begin{cases} 
\{ g : g^T(z - x) \leq 0, \forall z \in C \} & \text{if } x \in C \\
\emptyset & \text{if } x \notin C
\end{cases} \]
Some calculus

- \( f_i : \mathbb{R}^n \mapsto (-\infty, \infty], \ i = 1, \ldots, m, \) proper, convex functions

\[
F = f_1 + \cdots + f_m
\]

Assume \( \bigcap_{i=1}^m \text{rint}(\text{dom}(f_i)) \neq \emptyset \) then (as sets)

\[
\partial F(x) = \partial f_1(x) + \cdots + \partial f_m(x), \ \forall x
\]

- \( \mathcal{C} = \bigcap_{i=1}^m \mathcal{C}_i, \) then \( \psi_{\mathcal{C}} = \psi_{\mathcal{C}_i} + \cdots + \psi_{\mathcal{C}_m}, \) so \( N_{\mathcal{C}} = N_{\mathcal{C}_1} + \cdots + N_{\mathcal{C}_m} \)

\( x^* \) solves \( \min_{x \in \mathcal{C}} f(x) \) \iff \( x^* \) solves \( \min_x (f + \psi_{\mathcal{C}})(x) \)

\[ \iff 0 \in \partial (f + \psi_{\mathcal{C}})(x^*) \iff 0 \in \nabla f(x^*) + N_{\mathcal{C}}(x^*) \]
Special cases and examples

- Normal cone is a cone
- $x \in \text{int}(C)$, then $N_C(x) = \{0\}$
- $C = \mathbb{R}^n$, then $N_C(x) = \{0\}$, $\forall x \in C$
Special cases and examples

- Normal cone is a cone
- \( x \in \text{int}(C) \), then \( N_C(x) = \{0\} \)
- \( C = \mathbb{R}^n \), then \( N_C(x) = \{0\} \), \( \forall x \in C \)

EXAMPLE: POLYHEDRAL CASE

- \( C = \{z : a_i^T z \leq b_i, i = 1, \ldots, m\} \)
- \( N_C(x) = \left\{ \sum_{i=1}^{m} \lambda_i a_i : 0 \leq b_i - a_i^T x \perp \lambda_i \geq 0 \right\} \)
- \( \perp \) makes product of items around it 0, i.e.

\[(b_i - a_i^T x)\lambda_i = 0, \ i = 1, \ldots, m\]
Combining: KKT conditions

Example: convex optimization first-order optimality condition:

\[ x^\ast \text{ solves } \min_{x \in \mathcal{C}} f(x) \iff 0 \in \nabla f(x^\ast) + \mathcal{N}_\mathcal{C}(x^\ast) \]

\[ \iff 0 = \nabla f(x^\ast) + y, \; y \in \mathcal{N}_\mathcal{C}(x^\ast) \]

\[ \iff 0 = \nabla f(x^\ast) + y, \; y = A^T \lambda, \]

\[ 0 \leq b - Ax^\ast \perp \lambda \geq 0 \]

\[ \iff 0 = \nabla f(x^\ast) + A^T \lambda, \]

\[ 0 \leq b - Ax^\ast \perp \lambda \geq 0 \]

More generally, if \( \mathcal{C} = \{ z : g(z) \leq 0 \} \), \( g \) convex, (with CQ)

\[ x^\ast \text{ solves } \min_{x \in \mathcal{C}} f(x) \iff 0 \in \nabla f(x^\ast) + \mathcal{N}_\mathcal{C}(x^\ast) \]

\[ \iff 0 = \nabla f(x^\ast) + \nabla g(x^\ast) \lambda, \]

\[ 0 \leq -g(x^\ast) \perp \lambda \geq 0 \]
Variational Inequality (replace $\nabla f(z)$ with $F(z)$)

- $F : \mathbb{R}^n \to \mathbb{R}^n$
- Ideally: $C \subseteq \mathbb{R}^n$ – constraint set; Often: $C \subseteq \mathbb{R}^n$ – simple bounds

$$\text{VI}(F, C) : \quad 0 \in F(z) + N_C(z)$$

- VI generalizes many problem classes
- Nonlinear Equations: $F(z) = 0$ set $C \equiv \mathbb{R}^n$
- Convex optimization: $F(z) = \nabla f(z)$
- For NCP: $0 \leq F(z) \perp z \geq 0$, set $C \equiv \mathbb{R}^n_+$
- For MCP (rectangular VI), set $C \equiv [l, u]^n$.
- For LP, set $F(z) \equiv \nabla f(z) = p$ and $C = \{z : Az = a, Hz \leq h\}$. 
VI: $0 \in F(z) + \mathcal{NC}(z)$

Many applications where $F$ is not the derivative of some $f$
Other applications of complementarity

- Economics: Walrasian equilibrium (supply equals demand), taxes and tariffs, computable general equilibria, option pricing (electricity market), airline overbooking
- Transportation: Wardropian equilibrium (shortest paths), selfish routing, dynamic traffic assignment
- Applied mathematics: Free boundary problems
- Engineering: Optimal control (ELQP)
- Mechanics: Structure design, contact problems (with friction)
- Geology: Earthquake propagation
Nash problems

\[
\min_{x,t} \theta(x, t) \quad \text{s.t.} \quad F(x, t) = 0, \quad t \in T
\]

- \(x\) represents “state”, prices, production levels, etc
- \(t\) represents “taxes/tariffs”, strategy or design variables
- Often: \(F(x, t) = 0 \iff x = x(t)\) (implicit function)
Nash problems

\[ \min_{x_1, t_1} \theta_1(x_1, t) \text{ s.t. } F(x_1, t) = 0, \ t = (t_1, t_2) \in T_1 \]

- \( x \) represents “state”, prices, production levels, etc
- \( t \) represents “taxes/tariffs”, strategy or design variables
- Agent 1 problem is parameterized by agent 2’s variable
Nash problems

\[
\begin{align*}
\min_{x_1, t_1} \theta_1(x_1, t) \text{ s.t. } F(x_1, t) &= 0, \ t = (t_1, t_2) \in T_1 \\
\min_{x_2, t_2} \theta_2(x_2, t) \text{ s.t. } F(x_2, t) &= 0, \ t = (t_1, t_2) \in T_2
\end{align*}
\]

- \( x \) represents “state”, prices, production levels, etc
- \( t \) represents “taxes/tariffs”, strategy or design variables
- Solution is pair \((x_1, t_1), (x_2, t_2)\) so that each agent cannot improve when other agents strategy remains fixed
Bimatrix Games: Golden Balls

- VI can be used to formulate many standard problem instances corresponding to special choices of $M$ and $C$.
- Nash game: two players have $I$ and $J$ pure strategies.
- $p$ and $q$ (strategy probabilities) belong to unit simplex $\triangle_I$ and $\triangle_J$ respectively.
- Payoff matrices $A \in \mathbb{R}^{J \times I}$ and $B \in \mathbb{R}^{I \times J}$, where $A_{j,i}$ is the profit received by the first player if strategy $i$ is selected by the first player and $j$ by the second, etc.
- The expected profit for the first and the second players are $q^T Ap$ and $p^T Bq$ respectively.
- A Nash equilibrium is reached by the pair of strategies $(p^*, q^*)$ if and only if

$$p^* \in \arg \min_{p \in \triangle_I} \langle A q^*, p \rangle \text{ and } q^* \in \arg \min_{q \in \triangle_J} \langle B^T p^*, q \rangle$$
Optimal Sanctions (Boehringer/F./Rutherford)

- Sanctions can be modeled using similar formulations used for tariff calculations.
- Model as a Nash equilibrium with players being countries (or a coalition of countries).
- Demonstrate the actual effects of different policy changes and the power of different economic instruments.
- GTAP global production/trade database: 113 countries, 57 goods, 5 factors.
- Coalition members strategically choose trade taxes to
  1. maximize their welfare (no regrets) or
  2. \textit{minimize} Russian welfare (with willingness to pay weights - WTP).
- Russia chooses trade taxes to maximize Russian welfare in response.
- Explore unilateral vs Nash setting, and limits on Russian response.
Nash problems

\[ 0 \in \nabla_{(x_1, t_1)} \theta_1(x_1, t) + N_{C_1(t_2)}(x_1, t_1) \]

\[ 0 \in \nabla_{(x_2, t_2)} \theta_1(x_2, t) + N_{C_2(t_1)}(x_2, t_2) \]

- $x$ represents “state”, prices, production levels, etc
- $t$ represents “taxes/tariffs”, strategy or design variables
- Solution is pair $(x_1, t_1), (x_2, t_2)$ so that each agent cannot improve when other agents strategy remains fixed
- Note: implicit function $x = x(t)$ results in $x_1 = x_2$ at solution
- In convex case, can replace optimization problem by its KKT conditions, leading to MCP that is not KKT of a single optimization
Resulting Nash equilibrium with different coalition costs - all have big impact on Russia
- Restricting instruments can change effects
- Russia export taxes fixed at benchmark level
- Restricting instruments can change effects
- Equal yield constraint on Russian trade taxes
Complementarity Problems via Graphs

- \( \mathcal{T} = N_\mathbb{R}_+ = (\mathbb{R}_+ \times \{0\}) \cup (\{0\} \times \mathbb{R}_-) \)
- \( \mathcal{T} \) is “monotone”

\[-y \in \mathcal{T}(z) \iff (z, -y) \in \mathcal{T} \iff 0 \leq y \perp z \geq 0\]

By approximating (smoothing) graph can generate interior point algorithms for example \( yz = \epsilon, y, z > 0 \)

\[0 \in F(z) + N_{\mathbb{R}_+^n}(z) \iff (z, -F(z)) \in \mathcal{T}^n \iff 0 \leq F(z) \perp z \geq 0\]
Operators and Graphs ($\mathcal{C} = [-1, 1], \mathcal{T} = \mathcal{N}_\mathcal{C}$)

\[ z_i = -1, -F_i(z) \leq 0 \text{ or } z_i \in (-1, 1), -F_i(z) = 0 \text{ or } z_i = 1, -F_i(z) \geq 0 \]

\[ \mathcal{T}(z) \]

\[ \mathcal{T}^{-1}(y) \]

\[ (\mathcal{I} + \mathcal{T})^{-1}(y) = P_T(y) \]

\( P_T(y) \) is the projection of \( y \) onto \([-1, 1]\)
Normal Map

- Suppose $\mathcal{T}$ is a maximal monotone operator

\[ 0 \in F(z) + \mathcal{T}(z) \quad (GE) \]

- Define $P_{\mathcal{T}} = (I + \mathcal{T})^{-1}$ (continuous, single-valued, non-expansive)

\[ 0 \in F(z) + \mathcal{T}(z) \iff z \in F(z) + I(z) + \mathcal{T}(z) \]
\[ \iff z - F(z) = x \text{ and } x \in (I + \mathcal{T})(z) \]
\[ \iff z - F(z) = x \text{ and } P_{\mathcal{T}}(x) = z \]
\[ \iff P_{\mathcal{T}}(x) - F(P_{\mathcal{T}}(x)) = x \]
\[ \iff 0 = F(P_{\mathcal{T}}(x)) + x - P_{\mathcal{T}}(x) \]

This is the so-called Normal Map Equation
Normal manifold = \{ F_i + N_{F_i} \}

(Relative) interiors of faces $F_i$

form partition of $C$
\[ C = \{ z | Bz \geq b \}, \quad N_C(z) = \{ B'v | v \leq 0, \; v_{\mathcal{I}(z)} = 0 \} \]
\[ C = \{ z \mid Bz \geq b \}, \quad N_C(z) = \{ B'v \mid v \leq 0, v_{\mathcal{I}(z)} = 0 \} \]
\[ C = \{ z | Bz \geq b \}, \quad F(z) = Mz + q \]
The PATH algorithm

- Start in cell that has interior (face is an extreme point)
- Move towards a zero of affine map in cell
- Update direction when hit boundary (pivot)
- Solves or determines infeasible if $M$ is copositive-plus on $\text{rec}(C)$
- Solves 2-person bimatrix games, 3-person games too, but these are nonlinear

But algorithm has exponential complexity (von Stengel et al)
Theorem

Suppose $C$ is a polyhedral convex set and $M$ is an $L$–matrix with respect to $\text{rec} C$ which is invertible on the lineality space of $C$. Then exactly one of the following occurs:

- PATHAVI solves (AVI)
- the following system has no solution

\[ Mz + q \in (\text{rec} C)^D, \quad z \in C. \] (1)

Corollary

If $M$ is copositive–plus with respect to $\text{rec} C$, then exactly one of the following occurs:

- PATHAVI solves (AVI)
- (1) has no solution

Note also that if $C$ is compact, then any matrix $M$ is an $L$–matrix with respect to $\text{rec} C$. So always solved.
Experimental results: AVI vs MCP

PATH is a solver for MCP (mixed complementarity problem).

- Run PathAVI over AVI formulation.
- Run PATH over AVI in MCP form (poorer theory as \( \text{recC} \) larger).
- Data generation
  - \( M \) is an \( n \times n \) symmetric positive definite/indefinite matrix.
  - \( A \) has \( m \) randomly generated bounded inequality constraints.

<table>
<thead>
<tr>
<th>((m, n))</th>
<th>PathAVI</th>
<th>PATH</th>
<th>% negative eigenvalues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>status</td>
<td># iterations</td>
<td>status</td>
</tr>
<tr>
<td>(180,60)</td>
<td>S</td>
<td>55</td>
<td>S</td>
</tr>
<tr>
<td>(180,60)</td>
<td>S</td>
<td>45</td>
<td>S</td>
</tr>
<tr>
<td>(180,60)</td>
<td>S</td>
<td>2</td>
<td>F</td>
</tr>
<tr>
<td>(180,60)</td>
<td>S</td>
<td>1</td>
<td>F</td>
</tr>
<tr>
<td>(360,120)</td>
<td>S</td>
<td>124</td>
<td>S</td>
</tr>
<tr>
<td>(360,120)</td>
<td>S</td>
<td>55</td>
<td>S</td>
</tr>
<tr>
<td>(360,120)</td>
<td>S</td>
<td>2</td>
<td>F</td>
</tr>
<tr>
<td>(360,120)</td>
<td>S</td>
<td>1</td>
<td>F</td>
</tr>
</tbody>
</table>
What is EMP?

Annotates existing equations/variables/models for modeler to provide/define additional structure

- equilibrium
- $vi$ (agents can solve $\min / \max / vi$)
- bilevel (reformulate as MPEC, or as SOCP)
- dualvar (use multipliers from one agent as variables for another)
- QS functions (both in objectives and constraints)

Currently available within GAMS (full license available to course participants until August 8, 2016 - contact me!)

- Some solution algorithms implemented in modeling system - limitations on size, decomposition and advanced algorithms
- QS extensions to Moreau-Yoshida regularization, compositions, composite optimization
Splitting Methods

- Suppose $\mathcal{T}$ is a maximal monotone operator
  \[
  0 \in F(z) + \mathcal{T}(z) \quad (GE)
  \]
- Can devise Newton methods (e.g. SQP) that treat $F$ via calculus and $\mathcal{T}$ via convex analysis
- Alternatively, can split $F(z) = A(z) + B(z)$ (and possibly $\mathcal{T}$ also) so we solve (GE) by solving a sequence of problems involving just
  \[
  \mathcal{T}_1(z) = A(z) \text{ and } \mathcal{T}_2(z) = B(z) + \mathcal{T}(z)
  \]
  where each of these is “simpler”
- Forward-Backward splitting (or ADMM):
  \[
  z^{k+1} = (I + c_k T_2)^{-1} (I - c_k T_1) (z^k),
  \]