Optimality conditions and complementarity, variational inequalities, operators and graphs

Michael C. Ferris

University of Wisconsin, Madison

Funded by DOE-MACS Grant with Argonne National Laboratory

Auckland, April 26, 2018
The PIES Model (Hogan) - Optimal Power Flow (OPF)

\[
\min_x c(x)
\]

s.t. \( Ax \geq q \) \hspace{2cm} \text{balance}

\( Bx = b, x \geq 0 \) \hspace{2cm} \text{technical constr}

Such multipliers (LMP's) are critical to operation of market.

Can try to solve the problem iteratively (shooting method):

\( \pi_{\text{new}} \in \text{multiplier}(\text{OPF}(d(\pi))) \)
The PIES Model (Hogan) - Optimal Power Flow (OPF)

\[
\begin{align*}
\min_{x} & \quad c(x) \\
\text{s.t.} & \quad Ax \geq d(\pi) \\
& \quad Bx = b, x \geq 0
\end{align*}
\]

- \( q = d(\pi) \): issue is that \( \pi \) is the multiplier on the “balance” constraint
- Such multipliers (LMP’s) are critical to operation of market
- Can try to solve the problem iteratively (shooting method):
  \[
  \pi^{new} \in \text{multiplier}(OPF(d(\pi)))
  \]
Alternative: Form KKT of QP, exposing $\pi$ to modeler

$$L(x, \mu, \lambda) = c(x) + \mu^T (d(\pi) - Ax) + \lambda^T (b - Bx)$$

$$0 \leq - \nabla_\mu L = Ax - d(\pi) \quad \bot \quad \mu \geq 0$$
$$0 = - \nabla_\lambda L = Bx - b \quad \bot \quad \lambda$$
$$0 \leq \nabla_x L = \nabla c(x) - A^T \mu - B^T \lambda \quad \bot \quad x \geq 0$$

- **empinfo:** dualvar $\pi$ balance
- **Fixed point:** replaces $\mu \equiv \pi$
Alternative: Form KKT of QP, exposing $\pi$ to modeler

\[
0 \leq Ax - d(\pi) \quad \perp \quad \pi \geq 0
\]
\[
0 = Bx - b \quad \perp \quad \lambda
\]
\[
0 \leq \nabla c(x) - A^T \pi - B^T \lambda \quad \perp \quad x \geq 0
\]

- empinfo: dualvar $\pi$ balance
- Fixed point: replaces $\mu \equiv \pi$
- LCP/MCP is then solvable using PATH

\[
z = \begin{bmatrix} \pi \\ \lambda \\ x \end{bmatrix}, \quad F(z) = \begin{bmatrix} A \\ -A^T & -B^T \end{bmatrix} z + \begin{bmatrix} -d(\pi) \\ -b \\ \nabla c(x) \end{bmatrix}
\]

- Existence, uniqueness, stability from variational analysis
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
  $f : \mathbb{R}^n \mapsto (-\infty, +\infty]$
  \[ \partial f(x) = \{ g : f(z) \geq f(x) + g^T (z - x), \forall z \}, \]
  the subdifferential of $f$ at $x$

If $f$ is differentiable and convex, then $\partial f(x) = \{ \nabla f(x) \}$
- e.g. $f(z) = \frac{1}{2} z^T Qz + p^T z$, then $\partial f(x) = \{ Qx + p \}$
- $x^*$ 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

\[ g \in \partial \psi_C(x) \]

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

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

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 \( \psi_{\mathcal{C}} = \psi_{\mathcal{C}_1} + \cdots + \psi_{\mathcal{C}_m} \), 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 $N_C(x) = \{0\}, \forall x \in C$

$C = \{z : a_i^T z \leq b_i, i = 1, \ldots, m\}$
- polyhedral
- $N_C(x) = \{ \sum_{i=1}^{m} \lambda_i a_i : 0 \leq b_i - a_i^T x \perp \lambda_i \geq 0 \}$
- $\perp$ 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:

\[ x^* \text{ solves } \min_{x \in C} f(x) \iff 0 \in \nabla f(x^*) + N_C(x^*) \]

\[ \iff 0 = \nabla f(x^*) + y, \ y \in N_C(x^*) \]

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

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

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

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

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

\[ x^* \text{ solves } \min_{x \in C} f(x) \iff 0 \in \nabla f(x^*) + N_C(x^*) \]

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

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

- $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$
- 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: $0 \in F(z) + N_C(z)$

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

Complementarity can model fixed points and disjunctions

- 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

Good solvers exist for large-scale instances of Complementarity Problems
Complementarity Problems via Graphs

\[ T = \mathcal{N}_{\mathbb{R}_+} = (\mathbb{R}_+ \times \{0\}) \cup (\{0\} \times \mathbb{R}_-) \]

\[ T \] is “monotone”

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

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

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

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

\[ \mathcal{T}(z) \quad \mathcal{T}^{-1}(y) \quad (\mathcal{I} + \mathcal{T})^{-1}(y) = P_{\mathcal{C}}(y) \]

\((\mathcal{I} + \mathcal{T})^{-1}(y)\) is the projection of \(y\) onto \([-1, 1]\): \(P_{\mathcal{C}}(y)\)
Generalized Equations

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

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

- Define $P_\mathcal{T} = (I + \mathcal{T})^{-1}$

- If $\mathcal{T}$ is polyhedral (graph of $\mathcal{T}$ is a finite union of convex polyhedral sets) then $P_\mathcal{T}$ is piecewise affine (continuous, single-valued, non-expansive)

\[ 0 \in F(z) + \mathcal{T}(z) \iff z \in F(z) + I(z) + \mathcal{T}(z) \]

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

Use in fixed point iterations (cf projected gradient methods)
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):

  $$z^{k+1} = (I + c_k T_2)^{-1} (I - c_k T_1) (z^k)$$
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
Conclusions

- Convexity separates easy optimization problems from hard ones
- Modern convex analysis extends linear programming to richer but still tractable settings
- Modern optimization within applications requires multiple model formats, computational tools and sophisticated solvers
- Variational inequalities and set valued analysis important tools for big data problems
- Modeling, optimization, statistics and computation embedded within the application domain is critical
- Many new settings available for deployment; need for more theoretic and algorithmic enhancements
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 $q^T A p$ and $p^T B q$ respectively.
- A Nash equilibrium is reached by the pair of strategies $(p^*, q^*)$ if and only if

$$p^* \in \arg\min_{p \in \triangle_I} \langle A q^*, p \rangle \quad \text{and} \quad q^* \in \arg\min_{q \in \triangle_J} \langle B^T p^*, q \rangle$$
Formulation using complementarity

The optimality conditions for the above problems are:

\[-A q^* \in N_{\triangle I}(p^*) \text{ and } -B^T p^* \in N_{\triangle J}(q^*)\]

Therefore the corresponding VI is affine and can be written as:

\[0 \in \begin{bmatrix} 0 & A \\ B^T & 0 \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} + N_{\triangle I \times \triangle J}(\begin{bmatrix} p \\ q \end{bmatrix}). \quad (1)\]