SINGLE-STAGE / MULTISTAGE STOCHASTIC VARIATIONAL INEQUALITIES

Terry Rockafellar
University of Washington, Seattle
University of Florida, Gainesville

joint work with Roger J-B Wets

SPCOM 2015
Adelaide, South Australia
8–12 February 2015

“Generalized Equations” / “Variational Inequalities”

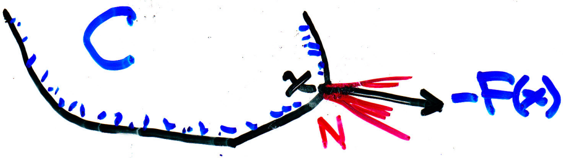
extending the classical paradigm of solving a system of equations

Variational inequality problem — in finite dimensions

For $C \subset \boldsymbol{R}^n$ nonempty closed convex, $F : \boldsymbol{R}^n \to \boldsymbol{R}^n$ continuous,

determine $x \in C$ such that $-F(x) \in N_C(x)$

i.e., $F(x)\cdot(x' - x) \geq 0 \ \forall x' \in C$



Modeling territory: optimality conditions, equilibrium conditions

Reduction to equation case: $N_C(x) = \{0\}$ when $x \in \mathrm{int}\, C$

$\Longrightarrow$ in case of $C = \boldsymbol{R}^n$, $\quad -F(x) \in N_C(x) \iff F(x) = 0$

Extending to a "Stochastic Environment"?

Underlying probability space: $(\Xi, \mathcal{A}, P)$

Problem elements subjected to uncertainty: $\xi \in \Xi$

- $C(\xi) \subset R^n$ closed convex $\neq \emptyset$, depending measurably on ξ
- $F(x,\xi) : R^n \times \Xi \to R^n$ continuous in x, measurable in ξ

BUT WHAT "PROBLEM" IS TO BE SOLVED?

Key question: which comes first, decision or observation?

Observation first: knowing ξ, respond by deciding $x(\xi)$

$-F(x(\xi),\xi) \in N_{C(\xi)}(x(\xi))$ a.s. a "random" V.I. problem?

Decision first: a single x must cope in advance with all ξ

$-F(x,\xi) \in N_{C(\xi)}(x)$ a.s. but is this hopeless to "solve"?

Conceptual limitation: anyway, why not more interaction?

maybe with information revealed and responded to in stages?

Review of Modeling Motivations for $-F(x) \in N_C(x)$

Elementary optimization: minimizing $g(x)$ over $x \in C$

$-\nabla g(x) \in N_C(x) \quad \longrightarrow$ first-order optimality, take $F = \nabla g$

Lagrangian V.I.: for $l(y,z)$ on $Y \times Z$ closed convex

$-\nabla_y l(y,z) \in N_Y(y), \quad \nabla_z l(y,z) \in N_Z(z),$ corresponding to

$x = (x,y), \ \ C = Y \times Z, \quad F(x) = (\nabla_y l(y,z), -\nabla_z l(y,z))$

$\longrightarrow$ this encompasses KKT conditions in NLP and much more!

Hierachical optimization/equilibrium:

- agent choosing $u \in U$ "controls" agenta(s) determining (y,z)
- minimization of $g(u,y,z)$ over $u \in U$ is desired

$-\nabla_u g(u,y,z) \in N_U(u), \ \ -(\nabla_y l(u,y,z), -\nabla_z l(u,y,z)) \in N_{Y \times Z}(y,z)$

$\longrightarrow$ modeled as a variational inequality in $x = (u,y,z)$ by taking:

$C = U \times Y \times Z, \quad F(x) = (\nabla_u g(u,y,z), \nabla_y l(u,y,z), -\nabla_z l(u,y,z))$

Back to Issues in S.V.I. Problem Formulation

Popular research focus: "solving" $-F(x,\xi) \in N_{C(\xi)}(x)$ a.s.
like finding a common solution to many optimization problems!

Fallback approach 1: "take expectations on both sides"

solve $-E_\xi[\,F(x,\xi)\,] \in N_D(x)$ for $D = \{\, x \mid x \in C(\xi) \text{ a.s.} \,\}$
solving a single V.I., but ad hoc? what interpretation?

Fallback approach 2: "find a best approximate solution"

minimize $E_\xi[\,f(x,\xi)\,]$ for some error or "gap" function f
not really "solving a V.I." and why useful to accomplish?

Imperatives for what a "stochastic variational inequality" should be

Formulations must be able to extend to a stochastic setting
the modeling capabilities of ordinary variational inequalities!

applications in stochastic programming? stochastic equilibrium?

Passing Instead to a Function Space Framework

Response function set-up

- Consider $x(\cdot) : \xi \mapsto x(\xi)$ in a space $\mathcal{L}^p_n = \mathcal{L}^p(\Xi, \mathcal{A}, P; \boldsymbol{R}^n)$

  pair $\mathcal{L}^p_n$ with $\mathcal{L}^q_n$, taking $\langle x(\cdot), v(\cdot) \rangle = E_\xi[\langle x(\xi), v(\xi) \rangle]$
- Introduce the closed convex set

$$\mathcal{C} = \{ x(\cdot) \in \mathcal{L}^p_n \mid x(\xi) \in C(\xi) \text{ a.s.} \}$$

- Introduce $\mathcal{F}$ as taking $x(\cdot) \in \mathcal{L}^p_n$ to an element $\mathcal{F}(x(\cdot)) \in \mathcal{L}^q_n$,

  $\mathcal{F}(x(\cdot)) : \xi \mapsto F(x(\xi)), \xi)$ maybe under more assumptions

Important formula to record:

$$-\mathcal{F}(x(\cdot)) \in N_{\mathcal{C}}(x(\cdot)) \iff -F(x(\xi), \xi) \in N_{C(\xi)}(x(\xi)) \text{ a.s.}$$

but this true V.I. in $\mathcal{L}^p_n$ isn't what we really want to solve

The challenge: adapt somehow to $x(\xi)$ NOT depending on ξ

Constancy as a Function Space Constraint

Substitute V.I. to investigate?

$$-\mathcal{F}(x(\cdot)) \in N_{\mathcal{C}_{\rm const}}(x(\cdot)) \;\text{ for }\; \mathcal{C}_{\rm const} = \left\{ x(\cdot) \in \mathcal{C} \,\middle|\, x(\cdot) \equiv x \text{ const} \right\}$$

Insight from stochastic optimization: a likely formula is

$$N_{\mathcal{C}_{\rm const}}(x(\cdot)) = \left\{ v(\cdot) - w(\cdot) \,\middle|\, v(\cdot) \in N_{\mathcal{C}}(x(\cdot)),\; E_\xi[w(\xi)] = 0 \right\}$$

$w(\cdot) \in \mathcal{L}^q_n$ serves as a Lagrange multiplier for constancy!

Conjecture: $-\mathcal{F}(x(\cdot)) \in N_{\mathcal{C}_{\rm const}}(x(\cdot)) \iff$

$$\begin{cases} x(\cdot) \equiv x \text{ const} \;\text{ and }\; \exists w(\cdot) \in \mathcal{L}^q_n,\; E_\xi[w(\xi)] = 0, \\ \text{such that } \; -F(x,\xi) + w(\xi) \in N_{C(\xi)}(x) \;\text{ a.s.} \end{cases}$$

Example: if $C(\xi) \equiv D$, this is equivalent to $-E_\xi[F(x,\xi)] \in N_D(x)$!

Justification hurdle: a "constraint qualification" is needed, and that may require working in $\mathcal{L}^\infty_n$, BUT generally $\mathcal{L}^1_n \neq \mathcal{L}^{\infty *}_n$

however there's no trouble in the **finitely stochastic** case

Multistage Format

Pattern of "decisions" and "observations" in N stages:

$x_1,\ \xi_1,\ x_2,\ \xi_2, \ldots, x_N,\ \xi_N$ with $x_k \in \boldsymbol{R}^{n_k},\ \xi_k \in \Xi_k$

$x = (x_1, \ldots, x_N) \in \boldsymbol{R}^n, \quad \xi = (\xi_1, \ldots, \xi_N) \in \Xi = \Xi_1 \times \cdots \Xi_N$

Nonanticipativity constraint

x_k can respond to $\xi_1, \ldots, \xi_{k-1}$ but not to $\xi_k, \ldots, \xi_N$:

$x(\xi) = (x_1, x_2(\xi_1), x_3(\xi_1, \xi_2), \ldots, x_N(\xi_1, \xi_2, \ldots, \xi_{N-1}))$

Nonanticipativity subspace: $\mathcal{N} \subset \mathcal{L}^\infty_n$

$$\mathcal{N} = \{ x(\cdot) \,|\, x_k(\cdot) \text{ depends only on } \xi_1, \ldots, \xi_{k-1} \}$$

$\longrightarrow$ $x(\cdot)$ is nonanticipative $\iff x(\cdot) \in \mathcal{N}$

Martingale subspace: $\mathcal{M} \subset \mathcal{L}^1_n$

$$\mathcal{M} = \{ w(\cdot) \,|\, E_{\xi_k, \ldots, \xi_N}[w_k(\xi_1, \ldots, \xi_{k-1}, \xi_k \ldots, \xi_N)] = 0 \}$$

$\longrightarrow$ in particular $E_\xi[w_1(\xi)] = 0$ and $w_N(\xi) \equiv 0$

Complementarity: $\mathcal{M} = \mathcal{N}^\perp, \quad \mathcal{N} = \mathcal{M}^\perp$

Single-stage example: $\mathcal{N} \longleftrightarrow x(\cdot)\ \mathrm{const}, \quad \mathcal{M} \longleftrightarrow E[w(\cdot)] = 0$

Proposed S.V.I. Problem Formulation

Other model ingredients as before:

$$\mathcal{C} = \{x(\cdot)\,|\,x(\xi) \in C(\xi) \text{ a.s.}\}, \quad \mathcal{F}(x(\cdot)) : \xi \mapsto F(x(\xi)), \xi)$$

but with $\mathcal{C} \subset \mathcal{L}_n^\infty$, $\mathcal{F} : \mathcal{L}_n^\infty \to \mathcal{L}_n^1$, $\mathcal{C}_{\rm const}$ upgraded to $\mathcal{C} \cap \mathcal{N}$

Stochastic variational inequalities — fundamentally

Basic form: $-\mathcal{F}(x(\cdot)) \in N_{\mathcal{C}\cap\mathcal{N}}(x(\cdot))$ "expandable to": (?)

Extensive form: $x(\cdot) \in \mathcal{N}$ and $\exists w(\cdot) \in \mathcal{M}$ such that

$$-F(x(\xi), \xi) + w(\xi) \in N_{C(\xi)}(x(\xi)) \text{ a.s.}$$

or **equivalently** as a V.I. on $x(\cdot)$ and $w(\cdot)$ jointly:

$$-(\mathcal{F}(x(\cdot)) + w(\cdot), -x(\cdot)) \in N_{\mathcal{C}\times\mathcal{M}}(x(\cdot), w(\cdot))$$

Stochastic variational inequalities — more broadly

$-\mathcal{F}(x(\cdot)) \in N_{\mathcal{K}\cap\mathcal{N}}(x(\cdot))$ for a closed convex set $\mathcal{K} \subset \mathcal{C}$

along with "Lagrange muliplier elaborations" of this

Orientation: reducing such a V.I. to basic or extensive form

S.V.I. Basic Form Versus Extensive Form

Outlook on the relationship:

- In the extensive form, $w(\cdot)$ is a nonanticipativity multiplier
- Invoking a multiplier rule requires a constraint qualification
- Otherwise the two conditions on $x(\cdot)$ should be equivalent
- Equivalence corresponds to confirming that

$$N_{\mathcal{C}\cap\mathcal{N}}(x(\cdot)) = N_{\mathcal{C}}(x(\cdot)) + N_{\mathcal{N}}(x(\cdot)), \text{ using } N_{\mathcal{N}}(x(\cdot)) \equiv \mathcal{M}$$

The finitely stochastic case: $(\Xi, \mathcal{A}, P)$ with Ξ **finite**, $\mathcal{A} = 2^{\Xi}$

- $\mathcal{L}_n^{\infty}$, $\mathcal{L}_n^1$, finite-dimensional, both identifiable as one "$\mathcal{L}_n$"
- relative interiors can serve in constraint qualifications

The more general stochastic case:

- $\mathcal{L}_n^1 \subset \mathcal{L}_n^{\infty *}$, $\neq$, with $\mathcal{L}_n^{\infty *} \backslash \mathcal{L}_n^1$ consisting of "singular elements"
- Singular elements could spoil the calculation of $N_{\mathcal{C}\cap\mathcal{N}}(x(\cdot))$
- Some way must be found to confine normals to $\mathcal{L}_n^1$, not $\mathcal{L}_n^{\infty *}$
- It will come from a 1976 Rock./Wets paper in multistage S.P.

Equivalence Results, First Part

Review of technical assumptions: behind $\mathcal{C}$ and $\mathcal{F}$

- $C(\xi) \neq \emptyset$, closed, convex, depending measurably on ξ
- $F(x, \xi)$ continuous in x, measurable **and integrable** in ξ

the integrability ensures that $\mathcal{F}(x(\cdot)) \in \mathcal{L}^1_n$

Sufficiency Theorem

If $x(\cdot)$ solves the S.V.I. in **extensive** form in partnership with some $w(\cdot)$, then $x(\cdot)$ also solves the corresponding S.V.I. in **basic** form

Necessity Theorem for the Finitely Stochastic Case

Suppose that the following **constraint qualification** is satisfied:

$$\exists\, \hat{x}(\cdot) \in \mathcal{N} \text{ such that } \hat{x}(\xi) \in \operatorname{ri} C(\xi) \text{ a.s.}$$

In that case, if $x(\cdot)$ solves the S.V.I. in **basic** form then $x(\cdot)$ with some $w(\cdot)$ also solves the corresponding S.V.I. in **extensive** form

this relies on calculus rules of **finite-dimensional** convex analysis

Additional Assumptions for the General Stochastic Case

Constraint boundedness — for the mapping $C : \xi \mapsto C(\xi)$

$\exists\, \rho > 0$ such that $C(\xi) \subset \rho B$ a.s. $\quad (B =$ unit ball in $R^n)$

on the side, this guarantees $\mathcal{C} \neq \emptyset$ in $\mathcal{L}_n^\infty$

Induced constraints? the choice of x_k in stage k can respond only to $(x_1, \ldots, x_{k-1})$ and $(\xi_1, \ldots, \xi_{k-1})$, and hence is limited to

$$C^k(x_1, \ldots, x_{k-1}, \xi) = \{ x_k \mid \exists (x_{k+1}, \ldots x_N) \text{ such that } (x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots x_N) \in C(\xi) \}$$

If this depends on future $(\xi_k, \ldots, \xi_N)$ it is necessary to constrain x_k to the **essential intersection** with respect to such information

Constraint nonanticipativity — no "induced constraints"

$C^k(x_1, \ldots, x_{k-1}, \xi)$ does not depend on $(\xi_k, \ldots, \xi_N)$

Equivalence Results, Second Part

Necessity Theorem for the General Stochastic Case

Assume **constraint boundedness and nonanticipativity**, and suppose the following **constraint qualification** is satisfied:

$$\exists\, \hat{x}(\cdot) \in \mathcal{N},\ \varepsilon > 0 \text{ such that } \hat{x}(\xi) + \varepsilon B \subset \operatorname{int} C(\xi) \text{ a.s.}$$

In that case, if $x(\cdot)$ solves the S.V.I. in **basic** form, then $x(\cdot)$ with some $w(\cdot)$ also solves the corresponding S.V.I. in **extensive** form

Method of proof: $-\mathcal{F}(x(\cdot)) \in N_{\mathcal{C}\cap\mathcal{N}}(y(\cdot))$ says that

$$x(\cdot) \in \operatorname{argmin}_{y(\cdot)} \left\{ \langle \mathcal{F}(x(\cdot)), y(\cdot) \rangle \,\middle|\, y(\cdot) \in \mathcal{C} \cap \mathcal{N} \right\}$$

- Recall that $\langle \mathcal{F}(x(\cdot)), y(\cdot) \rangle = E_\xi[\, \langle F(x(\xi), \xi), y(\xi) \rangle \,]$
- Introduce $f(y, \xi) = \langle F(x(\xi), \xi), y \rangle + \delta_{C(\xi)}(y)$
- Thus translate $-\mathcal{F}(x(\cdot)) \in N_{\mathcal{C}\cap\mathcal{N}}(x(\cdot))$ into **multistage S.P.:**

$$x(\cdot) \in \operatorname{argmin} \left\{ E_\xi[f(y(\xi), \xi)] \,\middle|\, y(\cdot) \in \mathcal{N} \right\}$$

- Get $w(\cdot) \in \mathcal{M}$ from result of that subject in Rock./Wets [1976]

Lagrangian Representations of Constraint Normals

Basic constraint system:

$$x(\xi) \in C(\xi) \iff x(\xi) \in X \text{ and } G(x(\xi),\xi) \in D$$

for $X \in R^n$, $D \subset R^m$ closed convex and $G : R^n \times \Xi \to R^m$

Multiplier rule: when D is a **cone** with **polar** Y there can be under a **constraint qualification** a Lagrangian formula

$$v(\xi) \in N_{C(\xi)}(x(\xi)) \iff \exists\, y(\xi) \in Y \text{ such that}$$
$$v(\xi) - \langle y(\xi), \nabla_x G(x(\xi),\xi)\rangle \in N_X(x(\xi)), \quad G(x(\xi),\xi) \in N_Y(y(\xi))$$

Lagrangian S.V.I. representation

Then for $\mathcal{X} = \{x(\cdot) \mid x(\xi) \in X \text{ a.s.}\}$ and $\mathcal{Y} = \{y(\cdot) \mid y(\xi) \in Y \text{ a.s.}\}$ the V.I. $-\mathcal{F}(x(\cdot)) \in N_{\mathcal{C}\cap\mathcal{N}}(x(\cdot)$ becomes a V.I. in $(x(\cdot), y(\cdot))$:

$$-(\mathcal{F}(x(\cdot)) + \langle y(\cdot), \nabla_x G(x(\cdot),\cdot)\rangle, -G(x(\cdot),\cdot)) \in N_{(\mathcal{X}\cap\mathcal{N})\times\mathcal{Y}}(x(\cdot), y(\cdot))$$

this is actually an S.V.I. of **basic** type with "$y(\cdot) = x_{N+1}(\cdot)$"!

Constraints Added to Basic Constraints

Expectation constraints: define $\mathcal{K} \subset \mathcal{C} \subset \mathcal{L}_n^\infty$ by adding

$$E_\xi[g_i(x(\xi),\xi)] \begin{cases} \leq 0 \text{ for } i = 1,\dots,r, \\ = 0 \text{ for } i = r+1,\dots,m \end{cases}$$

and as S.V.I. consider instead $-\mathcal{F}(x(\cdot)) \in N_{\mathcal{K}\cap\mathcal{N}}(x(\cdot))$

Reduction tactic: introduce multipliers λ_i for these constraints

$$\lambda = (\lambda_1,\dots,\lambda_m) \in \Lambda = [0,\infty)^r \times (-\infty,\infty)^{m-r}$$

Multiplier rule: under a constraint qualification

$$v(\cdot) \in N_{\mathcal{K}}(x(\cdot)) \iff \exists\, \lambda \in \Lambda \text{ such that}$$
$$v(\cdot) - \sum_{i=1}^m \lambda_i \nabla_x g_i(x(\cdot),\cdot) \in N_{\mathcal{C}}(x(\cdot)) \text{ and } G(x(\cdot)) \in N_\Lambda(\lambda)$$

where $G(x(\cdot)) = E_\xi[\,(g_1(x(\xi),\xi),\dots,g_m(x(\xi),\xi))\,]$

Reduced version of $-\mathcal{F}(x(\cdot)) \in N_{\mathcal{K}\cap\mathcal{C}}(x(\cdot))$ in these circumstances

$$-(\mathcal{F}(x(\cdot)) + \sum_{i=1}^m \lambda_i \nabla_x g_i(x(\cdot),\cdot), -G(x(\cdot)) \in N_{(\mathcal{C}\cap\mathcal{N})\times\Lambda}(x(\cdot),\lambda)$$

this is actually an S.V.I. of **basic** type with "x_1 augmented by λ"!

Some References

[1] X.J. Chen, R.J-B Wets and Y. Zhang **(2012)** "Stochastic variational inequalities: residual minimization smoothing/sample average approximations," *SIAM J. Optimization* 22, 649–673.

[2] R.T. Rockafellar and R.J-B Wets **(1976)** "Nonanticipativity and L^1-martingales in stochastic optimization problems," *Mathematical Programming Study* 6, 170-187.

[3] R.T. Rockafellar, R. J-B Wets **(1998)** *Variational Analysis*

[4] A.L. Dontchev, R.T. Rockafellar **(2014)** *Implicit Functions and Solution Mappings* **second edition!**

[5] R.T. Rockafellar, R.J -B Wets **(2015?)** "Stochastic variational inequalities: single-stage to multistage" $\approx$ **finished**

website: www.math.washington.edu/∼rtr/mypage.html