Solution Mappings for Variational Problems

Wolfgang Stockinger

Seminar Numerik

November 24, 2015

Overview

Generalized Equations

A condition on x of the form

$$f(x) + F(x) \ni 0 \iff -f(x) \in F(x),$$

with $f : \mathbb{R}^n \to \mathbb{R}^m$ and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is called generalized equation.

Examples:

- The zero mapping: $F \equiv 0$
- $F \equiv N_C$

Normal Cones

Definition

For a convex set $C \subset \mathbb{R}^n$ and a point $x \in C$, a vector v is said to be normal to C at x if $\langle v, x' - x \rangle \leq 0$ for all $x' \in C$. The set of all such vectors v is called the normal cone to C at x and is denoted by $N_C(x)$. For $x \notin C$, $N_C(x)$ is taken to be the empty set.

Variational inequalities

Definition

For a function $f : \mathbb{R}^n \to \mathbb{R}^n$ and a closed convex set $C \subset \operatorname{dom} f$ the generalized equation

$$f(x) + N_C(x) \ni 0$$

is called the variational inequality for f and C.

This definition is equivalent to the expression

$$x \in C, \quad \langle f(x), x' - x \rangle \geq 0 \quad \forall x' \in C.$$

Projection Mapping

A relation between the normal cone mapping N_C and the Projection mapping P_C is given by the following equivalence:

$$v \in N_C(x) \iff P_C(x+v) = x$$

The variational inequality can actually be written as an equation, namely

$$f(x) + N_C(x) \ni 0 \iff P_C(x - f(x)) - x = 0.$$

Solutions to Variational Inequalities

Theorem (Solutions to Variational Inequalities)

For a function $f : \mathbb{R}^n \to \mathbb{R}^n$ and a nonempty, closed convex set $C \subset \operatorname{dom} f$ relative to which f is continuous, the set of solutions to the variational inequality is always closed. It is sure to be nonempty when C is bounded.

Polar Cone

Theorem

Let K be a closed, convex cone and let K^* be its polar, defined by

$$K^* = \{y \mid \langle x, y \rangle \leq 0 \quad \forall x \in K\}.$$

Then K^* is likewise a closed, convex cone, and its polar $(K^*)^*$ is in turn K. Furthermore the normal vectors to K and K^* are related by

$$y \in N_K(x) \iff x \in N_{K^*}(y) \iff x \in K, \quad y \in K^*, \quad \langle x, y \rangle = 0.$$

Normal Cone/Tangent Cone

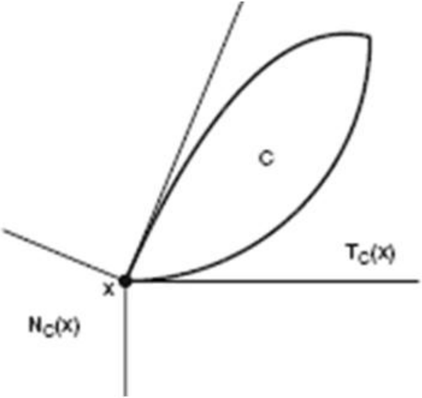

Figure: Normal and Tangent Cone to a convex set

Optimization Problems

For an objective function $g : \mathbb{R}^n \to \mathbb{R}$ and a constraint set $C \subset \mathbb{R}^n$ we consider a problem of the form

$$\text{minimize} \quad g(x), x \in C$$

Theorem (Basic Variational Inequality for Minimization)

Let $g : \mathbb{R}^n \to \mathbb{R}$ be differentiable on an open convex set O, and let C be a closed convex subset of O. In minimizing g over C, the variational inequality

$$\nabla g(x) + N_C(x) \ni 0,$$

is necessary for x to furnish a local minimum. It is both necessary and sufficient for a global minimum if g is convex.

Remarks

- When $C = \mathbb{R}^n$ we are dealing with unconstrained optimization
- The projection $x = P_C(z)$ is the solution to the minimization problem

$$\text{minimize } g(x) = \frac{1}{2}|x - z|^2$$

- If $C = C_1 \cap C_2$ for closed, convex sets C_1 and C_2 in $\mathbb{R}^n$, then the formula

$$N_C(x) = N_{C_1}(x) + N_{C_2}(x) = \{v_1 + v_2 \mid v_1 \in N_{C_1}(x), v_2 \in N_{C_2}(x)\}$$

holds for every $x \in C$ such that there is no $v \neq 0$ with $v \in N_{C_1}(x)$ and $-v \in N_{C_2}(x)$

Lagrangian Variational Inequalities

Theorem (Lagrange Multiplier Rule)

Let $X \subset \mathbb{R}^n$ and $D \subset \mathbb{R}^m$ be nonempty, closed, convex sets and consider the problem

$$\text{minimize} \quad g_0(x), \quad x \in C = \{x \in X \mid g(x) \in D\},$$

for $g(x) = (g_1(x), \ldots, g_m(x))$, where the functions $g_i : \mathbb{R}^n \to \mathbb{R}, i = 0, \ldots, m$ are continuously differentiable. Let x be a point of C at which the following constraint qualification condition is fulfilled:

$$\text{there is no } y \in N_D(g(x)), y \neq 0, \text{such that } -y\nabla g(x) \in N_X(x)$$

If g_0 has a local minimum relative to C at x, then there exists

$$y \in N_D(g(x)), \text{ such that } -[\nabla g_0(x) + y\nabla g(x)] \in N_X(x).$$

Remarks

- In the first order optimality condition y is said to be a Lagrange multiplier vector associated with x
- This condition can be reformulated by using the Lagrangian function, which is defined by

$$L(x, y) = g_0(x) + y_1 g_1(x) + \ldots + y_m g_m(x)$$

for $y = (y_1, \ldots, y_m)$.

Lagrangian Variational Inequalities

Theorem

In the previous minimization problem, suppose that the set D is a cone, and let Y be the polar cone D^*,

$$Y = \{y \mid \langle u, y \rangle, \quad \forall u \in D\}.$$

Then, in terms of the Lagrangian function, the condition on x and y can be written in the form

$$-\nabla_x L(x,y) \in N_X(x), \quad \nabla_y L(x,y) \in N_Y(y),$$

which furthermore can be identified with the variational inequality

$$(-\nabla_x L(x,y), \nabla_y L(x,y)) \in N_{X \times Y}(x,y).$$

The existence of $y \in Y$ satisfying this variational inequality with x is thus necessary for the local optimality of x.

Standard problem of nonlinear programming

Consider the problem

$$\text{minimize} \quad g_0(x), \quad g_i(x) \begin{cases} \leq 0, & \text{for } i \in [1, s] \\ = 0, & \text{for } i \in [s+1, m] \end{cases}$$

- $X = \mathbb{R}^n$
- D is the cone consisting of all $u = (u_1, \dots, u_m)$, such that $u_i \leq 0$ for $i \in [1, s]$ but $u_i = 0$ for $i \in [s+1, m]$.
- $D^* = Y = \mathbb{R}^s_+ \times \mathbb{R}^{m-s}$

The requirements for x and y are therefore

$$y \in \mathbb{R}^s_+ \times \mathbb{R}^{m-s}, \quad g_i(x) \begin{cases} \leq 0, & \text{for } i \in [1, s] \text{ for } y_i = 0 \\ = 0, & \text{for all other } i \in [1, m] \end{cases}$$

$$\nabla_x L(x, y) = 0.$$

Solution mapping

The object of study is now a parameterized generalized equation

$$f(p,x) + F(x) \ni 0$$

We consider the properties of the solution mapping $S : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ defined by

$$S : p \to \{x \mid f(p,x) + F(x) \ni 0\} \text{ for } p \in \mathbb{R}^d.$$

We first identify F with N_C, with $C \subset \mathbb{R}^n$ convex, closed and nonempty.

Theorem (Robinson Implicit Function Theorem)

For the solution mapping S to a parameterized variational inequality, consider a pair $(\bar{p}, \bar{x})$ with $\bar{x} \in S(\bar{p})$. Assume that:

- *$f(p, x)$ is differentiable with respect to x in a neighbourhood of the point $(\bar{p}, \bar{x})$, and both $f(p, x)$ and $\nabla_x f(p, x)$ depend continuously on (p, x) in this neighbourhood;*
- *the inverse G^{-1} of the set valued mapping $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ defined by*

$$G(x) = f(\bar{p}, \bar{x}) + \nabla_x f(\bar{p}, \bar{x})(x - \bar{x}) + N_C(x), \text{ with } G(\bar{x}) \ni 0,$$

  has a Lipschitz continuous single-valued localization σ around 0 for $\bar{x}$ with

$$lip(\sigma; 0) \leq \kappa.$$

Then S has a single-valued localization s around $\bar{p}$ for $\bar{x}$ which is continuous at $\bar{p}$, and moreover for every $\epsilon > 0$ there is a neighbourhood Q of $\bar{p}$ such that

$$\left|s(p') - s(p)\right| \leq (\kappa + \epsilon)\left|f(p', s(p)) - f(p, s(p))\right| \text{ for all } p', p \in Q.$$

Remarks

- If f is Lipschitz w.r.t. $p \Rightarrow s$ Lipschitz around $\bar{p}$
- If $C = \mathbb{R}^n \Rightarrow f(p, x) = 0$.
- The condition on G reduces in this case to the nonsingularity of the Jacobian $\nabla_x f(\bar{p}, \bar{x})$.

Theorem (Robinson Theorem Extended Beyond Differentiability)

For a generalized equation and its solution mapping S, let $\bar{p}$ and $\bar{x}$ be such that $\bar{x} \in S(\bar{p})$. Assume that:

- *$f(.,\bar{x})$ is continuous at $\bar{p}$, and h is a strict estimator of f with respect to x uniformly in p at $(\bar{p},\bar{x})$ with a constant μ;*
- *the inverse of G^{-1} of the mapping $G = h + F$, for which $G(\bar{x}) \ni 0$, has a Lipschitz continuous single-valued localization σ around 0 for $\bar{x}$ with $lip(\sigma;0) \leq \kappa$ for a constant κ such that $\kappa\mu < 1$.*

Then S has a single valued localization s around $\bar{p}$ for $\bar{x}$ which is continuous at $\bar{p}$, and moreover for every $\epsilon > 0$ there is a neighbourhood Q of $\bar{p}$ such that

$$\left|s(p') - s(p)\right| \leq \frac{\kappa + \epsilon}{1 - \kappa\mu} \left|f(p', s(p)) - f(p, s(p))\right| \text{ for all } p', p \in Q.$$

Theorem (Contracting Mapping Principle)

Consider a function $\varphi : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^m$ *and a point* $(\bar{p}, \bar{x}) \in \operatorname{int} \operatorname{dom} \varphi$ *and let the scalars* $\nu \geq 0, b \geq 0, a > 0$ *and the set* $Q \subset \mathbb{R}^d$ *be such that* $\bar{p} \in Q$ *and*

$$\left|\varphi(p, x') - \varphi(p, x)\right| \leq \nu \left|x - x'\right| \text{ for all } x', x \in \mathbb{B}_a(\bar{x}) \text{ and } p \in Q,$$
$$|\varphi(p, \bar{x}) - \varphi(\bar{p}, \bar{x})| \leq b \text{ for all } p \in Q.$$

Consider a set-valued mapping $M : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ *with* $(\bar{y}, \bar{x}) \in \operatorname{gph} M$ *where* $\bar{y} := \varphi(\bar{p}, \bar{x})$*, such that for each* $y \in \mathbb{B}_{\nu a + b}(\bar{y})$ *the set* $M(y) \cap \mathbb{B}_a(\bar{x})$ *consists of exactly one point, denoted by* $r(y)$*, and suppose that the function*

$$r : y \to M(y) \cap \mathbb{B}_a(\bar{x}) \text{ for } y \in \mathbb{B}_{\nu a + b}(\bar{y})$$

is Lipschitz continuous on $\mathbb{B}_{\nu a + b}(\bar{y})$ *with a Lipschitz constant* λ*. In addition suppose that*

- $\lambda \nu < 1$;
- $\lambda \nu a + \lambda b \leq a$

Theorem (Contracting Mapping Principle)

Then for each $p \in Q$ *the set* $\{x \in \mathbb{B}_a(\bar{x}) \mid x \in M(\varphi(p,x))\}$ *consists of exactly one point and the associated function*

$$s : p \to \{x \mid x = M(\varphi(p,x)) \cap \mathbb{B}_a(\bar{x})\} \text{ for } p \in Q$$

satisfies

$$\left|s(p') - s(p)\right| \leq \frac{\lambda}{1-\lambda\nu} \left|\varphi(p', s(p)) - \varphi(p, s(p))\right| \text{ for all } p', p \in Q.$$

Proof

Fix $p \in Q$ and define a function $\Phi_p : \mathbb{R}^n \to \mathbb{R}^m$ by

$$\Phi_p : x \to r(\varphi(p, x)) \text{ for } x \in \mathbb{B}_a(\bar{x}).$$

Let now $x, x' \in \mathbb{B}_a(\bar{x})$.

- Note that one has $|\bar{y} - \varphi(p, x)| \le b + \nu a$
- $|\Phi_p(\bar{x}) - \bar{x}| \le a(1 - \lambda\nu)$
- $|\Phi_p(x') - \Phi_p(x)| \le \lambda\nu\, |x' - x|$

$\Longrightarrow \Phi_p$ is Lipschitz continuous and has a unique fixed point $s(p)$ in $\mathbb{B}_a(\bar{x})$. Doing this for every $p \in Q$, we get a function $s : Q \to \mathbb{B}_a(\bar{x})$.

- $x = \Phi_p(x) \iff x = r(\varphi(p, x)) = M(\varphi(p, x)) \cap \mathbb{B}_a(\bar{x})$
- Since $s(p) = r(\varphi(p, s(p)))$ we get the estimate for $|s(p') - s(p)|$ by using triangular inequality, Lipschitz continuity of r and the estimates for φ.

Theorem

Let X be a Banach space and $\bar{x} \in X$. Consider a function $\Phi : X \to X$ for which there exist scalars $a > 0$ and $\lambda \in [0, 1)$ such that:

- $|\Phi(\bar{x}) - \bar{x}| \leq a(1 - \lambda)$;
- $|\Phi(x') - \Phi(x)| \leq \lambda\,|x' - x|$ *for every* $x, x' \in \mathbb{B}_a(\bar{x})$.

Then Φ has unique fixed point in $\mathbb{B}_a(\bar{x})$

Proof for the Robinson Theorem

For a fixed $\epsilon > 0$ choose $\lambda > lip(\sigma; 0), \nu > \mu$ such that $\lambda\nu < 1$

$$\frac{\lambda}{1-\lambda\nu} \leq \frac{\kappa+\epsilon}{1-\kappa\mu}$$

Choose positive numbers a, b, c such that

- $|\sigma(y) - \sigma(y')| \leq \lambda\,|y - y'|$ for $y, y' \in \mathbb{B}_{a\nu+b}(0)$
- $|e(p, x') - e(p, x)| \leq \nu\,|x - x'|$ for $x, x' \in \mathbb{B}_a(\bar{x})$ and $p \in \mathbb{B}_c(\bar{p})$
- $|f(p, \bar{x}) - f(\bar{p}, \bar{x})| \leq b$ for $p \in \mathbb{B}_c(\bar{p})$

Set $r = \sigma, M = (h+F)^{-1}, \bar{y} = 0$ and $\varphi = -e$. Furthermore observe that

$$x \in (h+F)^{-1}(-e(p,x)) \Longleftrightarrow x \in S(p).$$

By the contracting mapping principle we obtain that the solution mapping S has a single-valued localization around $\bar{p}$ for $\bar{x}$