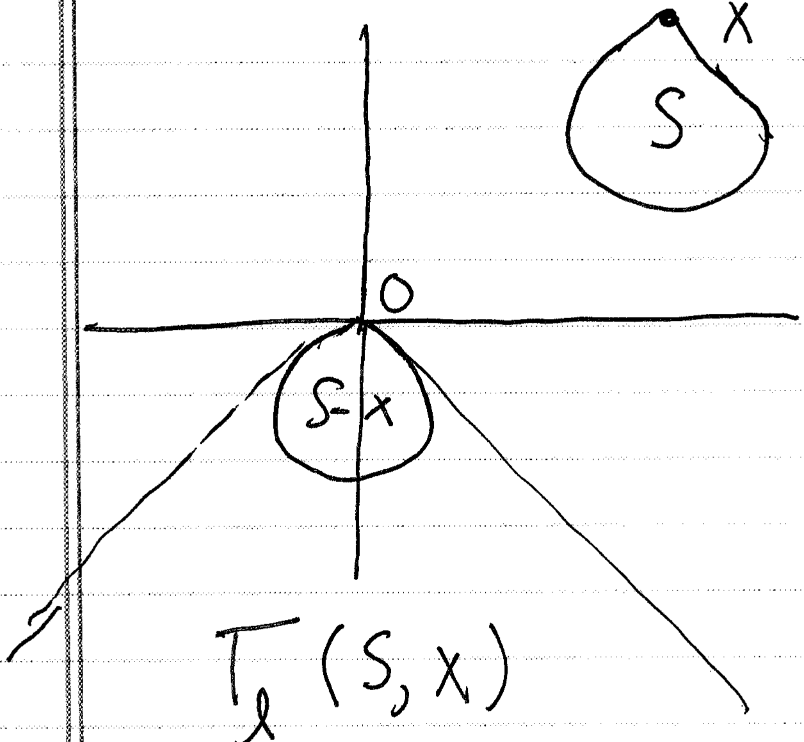
Chapter 7

Variational Inequality Necessity Theory and Applications

RECALL: FOR CONVEX S

$$\mathrm{CONE}(S-x) = T_\ell(S,x) \quad (1)$$



RECALL: FOR CONVEX $S \subset X$
X A NORMED LINEAR SPACE

$$\overline{T}(S,x) = T_c(S,x) = T_D(S,x)$$

and all are CONVEX.

Proposition 7.1.1. *Let S be a convex set in a real vector space X, and let $x \in S$. Then*

(i) *For any $y \in S$, we have that $y - x \in \cancel{T_S(x)}\ T_\ell(S,x)$.*

(ii) *For any $\eta \in \cancel{T_S(x)}\ T_\ell(S,x)$, there exist $\alpha > 0$ and $y \in S$ such that $\eta = \alpha(y - x)$.*

(iii) *Consequently,*

$$J'_+(x)(y-x) \geq 0 \ \ \textit{for all } y \in S \quad \Longleftrightarrow \quad J'_+(x)(\eta) \geq 0 \ \ \textit{for all } \eta \in \cancel{T_S(x)}\ T_\ell(S,x)$$

for any $J : S \to \mathbb{R}$ such that these derivatives exist.

4

Proposition 7.1.2. *Let S be a convex subset of a real vector space X, and suppose that $x^* \in S$. If $T_S(x^*)$ is symmetric, and if $J'_+(x^*)(\eta)$ is homogeneous in η for all $\eta \in T_S(x^*)$, then*

$$J'_+(x^*)(\eta) \geq 0 \text{ for all } \eta \in T_S(x^*) \iff J'_+(x^*)(\eta) = 0 \text{ for all } \eta \in T_S(x^*).$$

NOTE $\quad T_S(x^*) = T_l(S, x^*)$

5

The Variational Inequality

Theorem 7.1.3 (The Variational Inequality). *Consider $J : S \subset X \to \mathbb{R}$, where S is a convex subset of the real vector space X. Let x^* be a minimizer of J over S. Then, necessarily*

$$J'_+(x^*)(\eta) \geq 0 \quad \textit{for all } \eta \in T_S(x^*), \tag{7.1}$$

and equivalently,

$$J'_+(x^*)(y - x^*) \geq 0 \quad \textit{for all } y \in S, \tag{7.2}$$

whenever these one-sided derivatives exist. In addition, if J is convex on S, then (7.1), *equivalently* (7.2), *is also a sufficient condition for x^* to minimize J over S.*

Moreover, if $T_S(x)$ is symmetric and if $J'_+(x^*)(\eta)$ is homogeneous in η for all $\eta \in T_S(x)$, then the inequalitites in (7.1) *and* (7.2) *reduce to equalities.*

6

Proof. Suppose that x^* minimizes the function J over the convex set S, and let y be an element in S such that $J'_+(x^*)(y - x^*)$ exists. For the feasible arc emanating from x^*,

$$A(t) = x^* + t(y - x^*) \quad \text{for } t \in [0, 1], \quad \tau > 0,$$

define the function $\phi : [0, 1) \to I\!R$ by

$$\phi(t) \; = \; J\,(A(t)\,) \; = \; J(x^* + t(y - x^*)).$$

Then

$$\phi'_+(0) \; = \; J'_+(x^*)\,(y - x^*) \; \geq \; 0$$

by Theorem 3.3.3 (our fundamental principle for necessity), and we have necessity.

8

We now prove sufficiency. Suppose that J is convex on S. Furthermore, suppose that (7.2) holds at some point $x^* \in S$. Let y be any point in S. Then

$$\begin{aligned} 0 &\leq J'_+(x^*)(y-x^*) &&= \lim_{t\downarrow 0} \frac{J\left[x^* + t(y-x^*)\right] - J(x^*)}{t} \\ &\leq \lim_{t\downarrow 0} \frac{t\,J(y) + (1-t)\,J(x^*) - J(x^*)}{t} &&= J(y) - J(x^*). \end{aligned}$$

Hence x^* is a minimizer of J over S.

7.2 Applications of the Variational Inequality

In our applications of the variational inequality for the problem of minimizing $J : X \to \mathbb{R}$ over the set S contained in the real vector space X, we will follow the following five steps:

Step 1: Demonstrate that S is convex. If it is not, then the variational inequality does not apply to this problem.

Step 2: Calculate $\mathcal{T}_S(x)$ if it can be readily done; otherwise use the $\mathcal{D}_S(x)$ form of the variational inequality.

Step 3: Calculate $J'_+(x)(\eta)$ and $J''_+(x)(\eta, \eta)$.

Step 4: Examine $J''_+(x)(\eta, \eta)$ and check for convexity or strict convexity using $\mathcal{T}_S(x)$ or $\mathcal{D}_S(x)$ as appropiate.

Step 5: Find x^* satisfying the variational inequality or equality as appropriate.

As we shall soon see, the bulk of the work and need for creativity is in Step 5.

9

Helpful Observations

- The variational inequality in the form (7.2) is mathematically clean. It does not require the notion of the feasibility cone. However, the first form of the variational inequality, (7.1), stated in terms of the feasibility cone is often much easier to use. We have found this to be the case when the convex set S is defined by equalities and strict inequalities. However, in the case the definition of S involves nonstrict inequalities, and it is not known if the point under consideration strictly satisfies these inequalities, the second form of the variational inequality (7.2) is more convenient. The reader should understand why strict inequalities can be ignored in the calculation of the feasibility cone $T_s(x^*)$.

- If linear equalities are a part (perhaps all) of the defining relations for the constraint set S, then the feasibility cone, $T_s(x^*)$, will consist of the null spaces of the linear operators in these defining relations. To see this, observe that if $S = \{x \in X : Ax = b\}$, where A is linear, and $A(x^*) = b$, then $A(x^* + t\eta) = b$ reduces to $A(\eta) = 0$ since $t \neq 0$. Hence $T_s(x^*)$ is the null space of A.

- Suppose that we have found x^* a solution of the variational inequality ((7.1) or (7.2)). Now if J is strictly convex on S, then x^* is not only a solution, but is the unique solution of the optimization problem. Moreover, if J is convex, and x^* is the unique solution of the variational inequality, then x^* is again the unique solution of the optimization problem.

12

7.3 The Least-Squares Problem

We consider the linear system $Ax = b$, where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$. Our objective is to find an x that minimizes the Euclidean norm of the residuals. Equivalently, we solve the optimization problem

$$\underset{x}{\text{minimize}} \ \langle Ax - b, Ax - b \rangle, \tag{7.3}$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product in $\mathbb{R}^m$.

Theorem 7.3.1. *Least-squares solutions of the linear system $Ax = b$ exist. Moreover, x^* is a least-squares solution if and only if x^* satifies the normal equations*

$$A^T Ax = A^T b.$$

Finally, if A has full column rank, then

$$x^* = (A^T A)^{-1} A^T b$$

is the unique least-squares solution.

Proof: The proof follows the five-step program outlined in §7.2. (14)

Step 1 and 2: The problem is unconstrained; hence $S = I\!R^n$, and $\mathcal{T}_S(x)$ is also $I\!R^n$. So convexity of S and symmetry of $\mathcal{T}_S(x)$ are immediate.

Step 3: By differentiating

$$\phi(t) = J(x + t\eta)$$

twice with respect to t, we see that

$$J'(x)(\eta) = \langle A\eta, Ax - b\rangle \qquad \text{and} \qquad J''(x)(\eta,\eta) = \langle A\eta, A\eta\rangle \geq 0.$$

Step 4: It follows from our differential characterizations that J is convex.

Step 5: We have

$$J'(x)(\eta) = \langle \eta, A^T(Ax - b)\rangle.$$

Clearly $A^T(Ax - b) = 0$ implies $J'(x) = 0$. By choosing $\eta = A^T(Ax - b)$ for any given x, we see that $J'(x)(\eta) = 0$ only if $A^T(Ax - b) = 0$. This demonstrates that x^* is a least-squares solution if and only if x^* satisfies the normal equations. The fact that the normal equations have a solution and hence that least-squares solutions exist follows from linear algebra by way of Problem 2.5.2.

15

In the particular instance that A has full column rank, we know that $A^T A$ is positive definite. To see this, observe that $\eta^T A^T A\eta = \langle A\eta, A\eta\rangle \geq 0$. Equality would imply that $A\eta = 0$ for $\eta \neq 0$, which would contradict the fact that we have linearly independent columns. Hence, in this case we would have that J is strictly convex and the unique least-squares solution is given by $x^* = (A^T A)^{-1} A^T b$. This proves the theorem. $\square$