NOTE ON THE EQUIVALENCE OF KUHN-TUCKER COMPLEMENTARITY CONDITIONS TO AN EQUATION

Andrzej P. Wierzbicki

October 1981
WP-81-127

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

ABSTRACT

The note presents a more general and simple proof with geometric interpretations of the equivalence of the complementarity problem to an equation (or a system of equations) given by Mangasarian in 1976. Although this fact has been used by the author and others in a different context, it is believed that it should be presented to a more general audience of optimization specialists.

THE PROBLEM

Consider the optimization problem:

(1) minimize $f(x)$ subject to $g(x) \in -D$

where $f:E\to R^1$, $g:E\to F$, E,F are linear topological spaces, D is a closed convex cone in F. It is well known that, under additional smoothness and regularity assumptions, the necessary conditions for $\hat{x} \in E$, $\hat{\lambda} \in F^*$ being the primal and dual solutions of the problem (1) can be written as

(2) $f_x(\hat{x}) + g_x^*(\hat{x})\hat{\lambda} = 0$

(3) $g(\hat{x}) \in -D;\ \langle\hat{\lambda}, g(\hat{x})\rangle = 0;\ \hat{\lambda} \in D^*$

where F^* is the dual space to F, $\langle\cdot,\cdot\rangle$ is the duality relation between F and F^*, D^* is the dual cone to D, and $g_x^*(\hat{x})$ is the adjoint to $g_x(\hat{x})$, the Gateaux derivative of g at $\hat{x}$.

In the remaining part of this note, *we assume that* F *is a Hilbert space.* Thus F^* can be identified with F, $\langle\cdot,\cdot\rangle$ is the scalar product, and D^* is the polar cone to $-D$. If, furthermore, $F = R^m$ (with Euclidean norm, but this assumption is not essential in this special case) and $D = R^m_+$, then *the Kuhn-Tucker complementarity conditions (3)* can be written as

(4) $g(\hat{x}) \leq 0;\ \langle\hat{\lambda}, g(\hat{x})\rangle = 0;\ \hat{\lambda} \geq 0$

For this special case, one of Mangasarian's results (Ref.1) shows that (4) is equivalent to the following equation

(5) $(g(\hat{x}) + \hat{\lambda})_+ = \hat{\lambda}$

where $(\cdot)_+$ is the operation of taking positive part of a vector in R^m. However, the proof given by Mangasarian is algebraic and no geometric insight is given to this equivalence.

The equivalence of (4) and (5) has been actually used earlier by Rockafellar (Ref. 2) however, without specifying this as a separate result, only in the context of augmented Lagrangian functions, and also with algebraic proofs.

The purpose of this note is to present a simpler and more general proof of the equivalence (4) $\Longleftrightarrow$ (5), based on the geometrical interpretation illustrated in Figure 1.

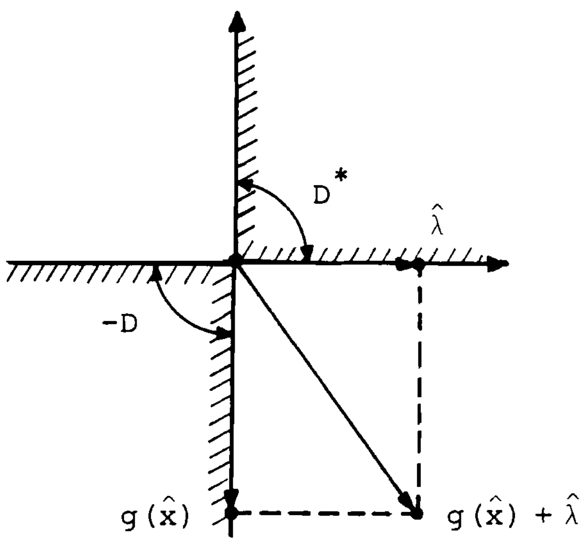

Figure 1. Geometrical interpretation of the equivalence of the equation $(g(\hat{x}) + \hat{\lambda})_+ = \hat{\lambda}$ to the Kuhn-Tucker complementarity condition.

The generalisation consists of the assumption that F is a Hilbert space and D is an arbitrary closed convex cone in F. Again, the result has been actually used by Wierzbicki and Kurcyusz (Ref. 3), however, only in the context of augmented Lagrangian functions for problems with constraints in a Hilbert space. Since then, the author has been persuaded[1] that the result has a value of its own, and should be known to a wider audience of optimization specialists, or even used when explaining seemingly complicated Kuhn-Tucker conditions to students. This is the main reason for publishing this note.

Theorem

Suppose F *is a Hilbert space,* $D \subset F$ *is a closed convex cone,* $\langle \cdot, \cdot \rangle$ *denotes the scalar product,* $D^* = \{y^* \in F^* = F: \langle y^*, y \rangle \geqslant 0 \text{ for all } y \in D\}$ *is the dual cone. Then the three following statments are equivalent to each other:*

(6) $g(\hat{x}) \in -D, \ \langle \hat{\lambda}, g(\hat{x}) \rangle = 0, \ \hat{\lambda} \in D^*$

(7) $(g(\hat{x}) + \hat{\lambda})^{D^*} = \hat{\lambda}$

(8) $(g(\hat{x}) + \hat{\lambda})^{-D} = g(\hat{x})$

where $(\cdot)^{D^*}$ *and* $(\cdot)^{-D}$ *denote the operations of projections on the cones* D^* *and* $-D$.

PROOF

The theorem is actually a corollary of the following theorem due to Moreau (Ref.4). Given a closed convex cone -D in a Hilbert space F and its polar cone D^*, any element $y \in F$ can be uniquely, orthogonally (and norm-minimally - see Wierzbicki and Kurcyusz (Ref.3)) decomposed into its projections on the cones -D and D^*.

---

(1) By many of his friends, but mostly by Terry Rockafellar and Olvi Mangasarian, to whom the author would like to express his thanks for encouragement.

In other words, Moreau's theorem reads: $y_1 = y^{-D}$ and $y_2 = y^{D*}$ are the projections of y on -D and D^* if, and only if

(9) $y_1 + y_2 = y;\ y_1 \in -D;\ y_2 \in D^*,\ \langle y_2, y_1 \rangle = 0$

Denote $g(\hat{x}) + \hat{\lambda} = y$, $g(\hat{x}) = y_1$, $\hat{\lambda} = y_2$. Then by Moreau's theorem, (6) implies (7) and (8). Suppose (7) holds. Then $y_2 = \hat{\lambda} = y^{D*}$. By Moreau's theorem, $y^{-D} = y - y^{D*} = g(\hat{x}) + \hat{\lambda} - \hat{\lambda} = g(\hat{x}) = y_1$ and (8) also holds. Conversely, (8) implies (7) by the same argument. But (7) and (8) together imply, by Moreau's theorem, that (6) holds. Thus, (6), (7) and (8) are mutually equivalent.

The theorem and its proof have clear geometrical interpretation as illustrated in Figure 2.

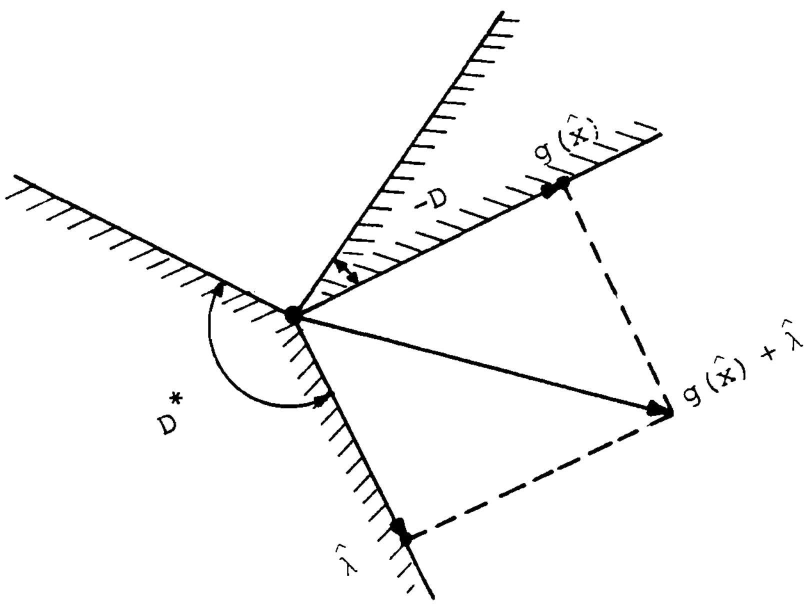

Figure 2. Geometrical interpretation of the equivalence of $(g(\hat{x}) \in -D,\ \hat{\lambda} \in D^*,\ \langle \hat{\lambda}, g(\hat{x}) \rangle = 0) \Longleftrightarrow (g(\hat{x}) + \hat{\lambda})^{D*} = \hat{\lambda} \Longleftrightarrow (g(\hat{x}) + \hat{\lambda})^{-D} = g(\hat{x})$.

COMMENTS

There are many possible implications and further properties of the equations equivalent to the Kuhn-Tucker complementarity conditions. They will be only outlined in these comments.

The equivalence (6) $\Leftrightarrow$ (7), taken together with (2), can be used to simplify sensitivity analysis of optimal solutions - since an implicit function theorem can be used to investigate the dependence of solutions of (2),(7) on possible parameters in the problem (1). The optimality conditions (2),(7) are equivalent to saddle-point conditions for an augmented Lagrangian function and have been exploited in this way. The conditions (2),(7) can be also used for a unification and a better understanding of many nonlinear programming algorithms. There are also many possible applications and interpretations in mathematical economics for equilibria described by complementarity conditions, etc.

Neither the condition (6) nor the equivalent conditions (7) or (8) define $\hat{\lambda}$ uniquely (first when taken together with (2), they might result in the uniqueness of $\hat{x},\hat{\lambda}$, under additional regularity assumptions). In fact, take any scalar $\varepsilon>0$ and substitute $\hat{\lambda}$ by $\varepsilon\hat{\lambda}$; this does not influence the validity nor equivalence of (6), (7),(8).

The operation of projection on a cone is not necessarily differentiable. If $F = R^m$ and $D = R^m_+$, then it is easy to show that the differentiability of $(g(\hat{x})+\hat{\lambda})^{D*} = (g(\hat{x})+\hat{\lambda})_+$ -- say, with respect to $\hat{\lambda}$ -- is equivalent to the fully complementarity: $(g(\hat{x})+\hat{\lambda})_+$ is differentiable if and only if there are no components $g_i(\hat{x}),\hat{\lambda}_i$ such that $g_i(\hat{x}) = 0$, $\hat{\lambda}_i = 0$. Thus, the left-hand sides of the system of equations (2),(5) can be differentiated only under full complementarity assumptions. However, if full complementarity does not hold, nondifferentiable analysis can be applied -- for example, the implicit function theorem for nondifferentiable mappings as given by Clarke, (Ref.5). In an infinite-dimensional case, the differentiability of a projection on a cone is a more complicated problem, but still preserves some similarity to full complementarity assumptions.

REFERENCES

1. Mangasarian, O.L. 1976. Equivalence of the complementarity problem to a system of nonlinear equations. SIAM Journ. Appl. Math. 31:89-92.

2. Rockafellar, R.T. 1976. Augmented Lagrangian multiplier functions and duality in nonconvex programming. SIAM Journ. Control. 12:268-285.

3. Wierzbicki, A.P., and St Kurcyusz. 1977. Projection on a cone, penalty functionals and duality theory for problems with inequality constraints in Hilbert space. SIAM Journ. Control and Optimization. 15:25-56.

4. Moreau, J.J. 1962. Décomposition orthogonale d'un espace Hilbertien selon deux cônes mutuellement polaires. C.R. Acad. Sci: Paris. 225:238-240.

5. Clarke, F.H. 1974. On the inverse function theorem. Pacific Journ. Math. 9:97-102.