EUROPEAN UNIVERSITY INSTITUTE, FLORENCE

DEPARTMENT OF ECONOMICS

E U I   W O R K I N G   P A P E R   No. 69

ON THE FORMALIZATION OF POLITICAL PREFERENCES : A CONTRIBUTION TO THE FRISCHIAN SCHEME

by

Berc RUSTEM and Kumaraswamy VELUPILLAI

October 1983

Berc Rustem
Programme of Research in
Optimal Policy Evaluation
Imperial College of Science and Technology
University of London

Kumaraswamy Velupillai
Department of Economics
European University
Institute
S. Domenico di Fiesole

"In mathematics EVERYTHING is algorithm and NOTHING is meaning; even when it doesn't look like that because we seem to be using WORDS to talk ABOUT mathematical things. Even these words are used to construct an algorithm."

Ludwig Wittgenstein
(in Philosophical Grammar, Ch. VII, §40)

BADIA FIESOLANA, SAN DOMENICO (Fi)

Printed in Italy in October 1983
European University Institute
Badia Fiesolana
50016 San Domenico (Fi)
Italy

1. Introduction *

The somewhat uncritical appeal to 'Arrow Impossibilities' in rejecting any notion of aggregate criterion functions to provide more structure to problems in the optimal formulation of economic policy seems, almost always, due to confusing Arrow's aims with the general problem of preference in relation to value theory in general, and to its three principal branches — aesthetics, economics and ethics. Recent work by Sen (1973), Harsanyi (1977, esp. Ch.4), Johansen (1969) and others, indicates that serious attempts, at a theoretical level, are being made to delineate more clearly the precise domain of relevance of impossibility theorems in relation to concepts of pure preference (cf. von Wright, 1963). As Johansen very pungently noted:

> "I think the existence question [of an aggregate objective function] is wrongly put. *It exists if we have constructed it, and the relevant question is whether we will find it useful and practicable to establish it*". (Johansen, 1979, p.108: italics added).

Johansen was, perhaps, only restating more concisely the almost passionate appeals made by that great initiator and practitioner (and indeed Johansen's own teacher) of quantitative economic policy: Ragnar Frisch. In his Nobel Prize Lecture - recently reprinted in a special volume of the American Economic Association (cf. Frisch, 1981) - and even more emphatically in his highly detailed last-published work[1] (cf. Frisch, 1972) Frisch made a strong plea for a 'Cooperation Between Politicians and Econometricians on the Formalization of Political Preferences'. The theory of economic policy owes a great deal for its quantitative developments to the pioneering works of Frisch, Tinbergen, Theil and Bent Hansen

* An earlier version of this paper was presented at a seminar in the Faculty of Economics and Politics of the University of Bologna.

(cf. Tinbergen, 1983, pp.161-162). However, in spite of early work by Frisch himself (Frisch, 1956; 1959; 1961) the delicate problem of formalizing the (aggregate) criterion (or preference) function traditionally employed in deriving optimal economic policy rules seems to have been bedevilled by the above-mentioned confusion with Arrow's important results. Frisch, in referring to objectives based on Arrow 'type' results to the use of an aggregate *preference function* 'which must underlie the very concept of an *optimal* economic policy', went on to observe (and here we choose to quote extensively so as to avoid misunderstandings about the precise nature of our own aims):

> "It is said that there are *many different* systems of preferences. It is impossible to choose between these systems. Therefore the concept of a preference function cannot be used in connection with national models. *This is one of the biggest pitfalls in this discussion of this matter.* Of course, there are differences of opinion. One social group may have one type of preference and another social group may have other preferences, and different persons may have different preferences, and even the same person may have different preferences at different points of time. All this is, of course, true. But the problem of settling differences of opinion *is not a special problem of econometrics.* It is a general problem of human behaviour and opinions. And there exists a machinery for settling such differences. This machinery is simply the political system of the country. This political system - whatever it may be - has been created for the very purpose of settling such differences. *What we have to do as econometricians is to apply this very system for the formalization of the preferences to go with our models.* Thus the preference function as it appears in our models is an expression of the preferences of the decision-making authority, whatever that authority may be. *The preference function in the model must not be confused with a general 'Welfare function' in the sense of welfare theory*".
> (Frisch, 1972, pp.7-8: italics added, and cf. also, Frisch, 1961, p.44 and f.n.2, p.44).

Interpreting Frisch's strictures almost literally, we try, in this paper, to provide an iterative scheme to enable a policy maker and an econometrician (model builder) to cooperate in the formalization of political preferences in the sense of trying to quantitatively clarify what the political authorities really are aiming at.

Thus, in Section 2, a summary of Frisch's formulation of the problem is given. In Section 3, on the basis of our interpretation of Frisch's formulation, a formalized positive, constructive, solution in the form of an iterative scheme to encapsulate Frisch's 'interview technique' is offered. In Section 4 some technical remarks on computational considerations and convergence problems are discussed.
In Sections 5 and 6 we discuss extensions of our method, again within the aims set forth by Frisch in the above two publications and, in particular, to the problems of the 'Optimal Price of a Bound' (Frisch, 1972; pp.14-17), i.e., shadow prices. Finally, in Section 7 we discuss the complexity of the policy design process implied by the method, using Khatchian's (1979,1980) ellipsoid algorithm.

2. Frisch on the "Cooperation Between Politicians and Econometricians on the Formalization of Political Preferences"

Leif Johansen's survey (Johansen, 1974) provides a comprehensive and highly readable account of Ragnar Frisch's contributions towards solving the problem of constructing *political preference functions* (cf. Frisch, 1956; p.45 ff and Frisch, 1961; p.44), for so-called *macroeconomic decision models* (or for the purpose of *macroeconomic programming*). We need, therefore, to concentrate only on a summary of the essentials of Frisch's approach and refer the interested reader to Johansen's detailed article.

In a series of articles, stretching over a period of over forty years, Ragnar Frisch was concerned with the problem of *constructing* what he called political preference functions for use in deriving optimal macroeconomic policies. The econometrician's duty (cf. below, f.n.2), he maintained, was to make it possible for the political decision maker(s) to be aware of what the economy was able to do[3] so that they, in turn, can formulate their wishes regarding what it ought to do - or should be made to do such that policies can be devised in an efficient way to achieve the latter desires. The policy maker[4] was expected to fix certain bounds for the levels and rates of change of politically sensitive variables in addition to the coefficients of a preference function indicating relative desires between (those bounded variables themselves and any) other necessary additional variables[5]. Frisch very clearly pointed out that *the policy maker's awareness of what was feasible determined, to a large extent*, the nature and scope of the bounds (s)he would tend to consider, as well as the relative weights that were to be attached to relevant variables in the preference function. It was therefore necessary, he argued, that the econometrician and the policy maker cooperate in such a way that an increasing 'perceptibility of the variable from the preference view point' (Frisch 1961, p.47) and a clearer indication of the nature of the bounds that *must* be imposed will become possible. Together with an element of consistency on the part of the policy maker - with the dose of realism or pragmatism that this class of worthy persons seem naturally to possess (an 'overdose' some idealists would contend!) - this cooperation with the econometricians should, iteratively, lead to a set of bounds and weights (for the preference functions) such that a politically acceptable set of values for the variables of interest can be derived using an optimal set of economic policies. Put another way, the idea is that the economet- rician, as a neutral expert, should be able to advise on the optimality of economic policies that can generate a politically acceptable set of trajectories such that the interlocking nature of desires, constraints and weights will be made very clear to the policy maker who, as a result of this, reveals, almost unwittingly, his desires and dreams whilst, hopefully, shedding some, at least, of his delusions.

The particular technique of such a cooperation between a policy maker and an econometrician was to be a consecutive series of carefully constructed interviews. Frisch, in the papers already cited above, investigated in great detail the efficient formulation of interview techniques. Basically there were three types of interview questions:

a) Questions such that alternatives would be ranked.

b) Questions of a dichotomic nature.

c) Distribution questions.

These three types of questions were related to the type of preference functions he was trying to construct. The first type of question was devised with the aim of establishing linear preference functions whereas the latter two were used, in combination, in his attempts to construct quadratic preference functions.

The sequence of questions, and hence the envisaged cooperation between econometricians and policy maker, itself was to take the following form:

a) In a preparatory phase, the econometrician, armed with a comprehensive model of the economy and 'making a systematic use of his general knowledge of the political atmosphere in the country, and in particular the political atmosphere in the party in question to which a constructed preference function would apply' (Frisch 1972; p.10), should form 'a tentative opinion' about the weights to choose.

b) Using the above weights, and depending on the type of preference functions to be constructed (i.e., linear, quadratic, cubic, etc.) optimal values for relevant variables are generated and the policy maker is confronted with them, in terms of one of the above three types of questions.

c) The answers, which cannot be arbitrary given the carefully structured questions, will generally indicate the nature of changes needed 'in the formulation of the preference function in order to produce a solution that comes closer to what the politicians .... say they want' (Frisch, 1972; p.12).

d) The task of the econometrician, at this stage, is to translate these vague indications to concrete changes in the weights of the preference function, such that a new solution closer to what the policy maker desires can be found.

e) Using these 'new' weights a new set of optimal solutions[6] will be - should be - generated by the econometrician (and his collaborators). These new solutions, closer to what the policy makers say they wanted, when next presented to them may elicit the following response:

> "No, this was not really what we wanted .... we have to change these particular aspects of your solution" (Frisch 1972, p.12).

f) Earlier steps are then repeated with the result (hopefully!) that:

> "[it] .... leads to a discussion back and forth. In this way one will work *step by step towards* a preference formulation such that the politicians can say about the resulting solution: 'All right, this is what we should like to see'. Or perhaps the expert will have to end by saying politely: 'Your Excellencies, I am sorry but you cannot, at the same time, have all these things on which you insist'. Their excellencies, being intelligent persons, will understand the philosophy of the preference questions .... and will .... acquiesce in a solution which is not quite what they like, but at least something better than other alternative lines of the development course which have emerged from the previous tentative formulations of the preference function". (Frisch, 1972; p.13: italics added.)

At this point, of course, the iteration ends. However, the following additional points, reflecting various observations made by Frisch at various stages over four decades of grappling with this problem, must be noted.

(i) Though Frisch did not consider anything other than linear, quadratic and cubic preference functions - and thus the interview questions were formulated with such functions in mind - he was aware of technical devices whereby more general functional forms could be reduced to the above simpler forms. In fact he worked out the details only for the case of a separable quadratic preference function. (But, cf. Frisch, 1972 , Ch.7, pp.30-32.) However, in discussing the problem of 'the bounds' (or the constraints) he did note that, in some important cases, general nonlinearities in the preference function can be 'transferred' to inequality constraints of the model subject to which the macroeconomic programming problem is solved.

(ii) Though we have, above, summarized an iterative scheme for the construction of *a* policy maker's preference function, it must be noted that Frisch went beyond the case of a single decision maker. But the details were not worked out as in the case for a single policy maker (cf., however, Frisch , 1959, 1967 , and our own attempts, in this spirit, in Rustem and Velupillai, 1979 , and Goodwin and Velupillai, 1982).

(iii) The importance of specifying correctly *the desired trajectories* was stressed by Frisch in almost all the above-cited articles. He did not assume that the policy maker would have, a priori, a 'correct' view of the desired trajectories. During the course of the iterative process of interview, results, re-evaluations the policy maker was supposed to become clearer, not only about what he wanted, but also about the feasible set of the model. This, he maintained, should imply that the policy maker should force the model to work at capacity (cf. also Kornai , 1975, esp. p.420, ff.).

(iv) Frisch took great care to point out that the econometrician should devise techniques such that the policy maker will only have to respond with respect to variables of *direct* relevance - the rest should be taken care of by the model at the disposal of the econometrician. This point can best be exemplified by Johansen's following clear observation:

> "It may be necessary for the [econometrician] to explain .... more carefully to the politician and perhaps advise him on those aspects of the [indirect variables, e.g.:] 'preference for investment' which are not a question about [direct variables, or] pure preferences but also a question about the likely effects of [indirect variables on direct ones, i.e.:] investment on future consumption possibilities. ... [Thus,] .... the [econometrician] should take great care to explain to the politician that he should not think of investment as a means of stimulating income by a multiplier effect: *this effect should be taken care of by the model, and not interefere with the specification of the preferences*". (Johansen ,1974 , p.48: italics added.)

Naturally, it will be impossible to do full justice to the unique contributions made by Ragnar Frisch within the compass of a potted summary that we have attempted in this Section. However, the problem is important and so few seem to have been seriously interested in it (in spite of the almost indiscriminate use of optimal control and mathematical programming techniques at all levels of planning and analysis) that, even at the risk of some distortion, it may be worthwhile[7].

The formalization as an iterative process of Frisch's 'interview approach' towards constructing a preference function of a policy maker we outline in the next sections[8].

3. An Iterative Method to Formalize Frisch's Scheme

The problem we wish to formalize, then, is the case in which a policy maker desires to optimize a certain set of target variables which are, in general, functions of the remaining (e.g. decision) variables, subject to a set of constraints on all the variables[9]. The policy maker's implicit preference function is assumed, in turn, to be a quadratic function of the target variables. The existence of nonlinear relationships between the target and other variables, also noted by Frisch (c.f. (i), Section 2), enables one to consider more general nonlinearities (than the quadratic case in the preference function by replacing them with suitable relations in the constraint set. The essential point of Frisch's problem is that the policy maker's preference function is not known explicitly either to himself or the econometrician. It is to be elicited by means of a series of interactions between the policy maker and the econometrician - the latter equipped with the set of relations describing the feasible set of the economy and some intuition about the possible weights between the target variables. From this starting point an iterative sequence of optimization, reflection, re-evaluation between the policy maker and the econometrician should lead to a converging set of weights between the target variables as awareness develops of the nature of the economy's feasible set and the policy maker's (implicit) system of values.

Frisch's own positive solution - and other related approaches - rely heavily on some variant of the interview technique in that at each iteration, starting from an arbitrary (but informed) set of weights, <u>in the interaction</u> a class of values for the target set of variables is optimally generated and presented to the policy maker. The latter is then asked to rank the alternative sets in terms of desirability. The next iteration, based upon the weights underlying the most desirable alteratives, proceeds in a similar fashion. This exercise is repeated until, hopefully, some such point as depicted in (f) in Section 2 is reached.

There are, however, some undesirable features in the above procedure - and Frisch was well aware of them. This was why he devoted so much time to the obviously unenviable task of devising suitable questions. Quite apart from the difficult question of convergence, this procedure does not eliminate the need to ask the policy maker to rank alternatives in terms of desirability. It was precisely to avoid this, on the basis of his practical experience in India and Norway, that Frisch went to great pains in perfecting his interview techniques. Frisch repeatedly stressed the point that it was impossible, and indeed unfair, to expect policy makers to be explicit about desired rankings from a set of Pareto efficient alternatives.

Thus, in attempting to formalize the Frischian scheme we have paid particular attention to the problem of the convergnece of weights in a finite number of iterations and, more importantly, to avoid the need to force the policy maker to be explicit about ranking a set of Pareto efficient alternatives. Indeed, in our method, the policy maker is not confronted, at each iteration, with a set of Pareto efficient alternatives. He is, in fact, presented with one optimally generated set of values for the target variables and asked only to indicate preferred directions for each one of them. These preferred directions are then translated into corrections of the weights.

Let, therefore, $\underline{x} \varepsilon E^n$ be the vector of target variables and R be the feasible set of $\underline{x}$. Let the elements of $\underline{x}^d$ be the desired values of the corresponding variables in $\underline{x}$. The policy maker would ideally like to achieve $\underline{x} = \underline{x}^d$. Indeed, given the restriction R, a particular element or set of elements of $\underline{x}$ may attain their corresponding desired values if the rest of the elements are allowed to assume any value required to attain these desired values. However, usually some elements of $\underline{x}^d$ imply a somewhat conflicting desire with respect to other elements of $\underline{x}^d$. Thus, all the elements of $\underline{x}^d$ cannot be attained simultaneously and hence $\underline{x}^d$ is infeasible, i.e., $\underline{x}^d \notin R$. Clearly, when $\underline{x}^d \in R$, all the $\underline{x}$ values desired by the policy maker can be attained and thus the optimal policy would have the obvious solution $\underline{x} = \underline{x}^d$. This paper is therefore obviously concerned with the problems arising when $\underline{x}^d \notin R$. In such cases, the relative importance of each element of $\underline{x}$ attaining its desired value has to be determined. This information can then be used to compute the best feasible alternative on R to $\underline{x}^d$. In mathematical terms, this problem can be summarized as

$$\min\{q_c(\underline{x}) \mid \underline{x} \in R\} \tag{3.1}$$

where[10]

$$q_c(\underline{x}) \triangleq \tfrac{1}{2} \langle \underline{x} - \underline{x}^d, Q_c(\underline{x} - \underline{x}^d) \rangle \tag{3.2}$$

$$= \tfrac{1}{2} \| \underline{x} - \underline{x}^d \|^2_{Q_c} \tag{3.3}$$

and Q_c is a symmetric[11], positive semi-definite weighting matrix which specifies the relative importance of each element of $\underline{x}$ attaining its corresponding desired value. Clearly, by specifying a positive definite Q_c, a measure of distance from $\underline{x}$ to $\underline{x}^d$ is defined (see (3.3)). The iterative method of this paper is aimed at the tailoring of Q_c to meet the requirements of the policy maker regarding the target variables. The method is not concerned with a "best" set of weights *independent* of the desired value $\underline{x}^d$. A solution, optimally generated via (3.1), acceptable to the policy maker is the main aim. Once the initial optimal solution is determined using an initial Q_c, the method provides a systematic way in which the policy maker can specify his dissatisfaction with the various elements of the optimal solution and leaves it to the method to alter Q_c to generate a more acceptable optimal solution. The method translates the policy maker's dissatisfaction with the initial optimal solution into a rank-one correction to Q_c. It is proposed that an "optimal" weighting matrix will be obtained by repeated updating so that, at the end of the iterative procedure, the "final" optimal solution will be totally acceptable to the decision maker.

When Q_c is positive definite and R is convex, the solution of (3.1) has a simple geometric interpretation. As (3.1) minimizes the norm defined in (3.3), the solution is the projection of $\underline{x}^d$ on to R with respect to the norm (3.3) (see Luenberger, 1969; Rustem, 1981). When R denotes a set of linear equality constraints, e.e.

$$R = \{ \underline{x} \in E^n \mid N^T \underline{x} = \underline{b} \} \tag{3.4}$$

where $\underline{b}$ is an (m x 1) constant vector and N is an (n x m) matrix whose columns are assumed (for simplicity) to be linearly independent, this projection, and hence the solution of (3.1), is stated in the following Lemma.

Lemma 1

When Q_c is positive definite and R is given by (3.4), the solution of (3.1) is

$$\underline{x}_c = \underline{x}^d - H_c N (N^T H_c N)^{-1} (N^T \underline{x}^d - \underline{b}) \tag{3.5}$$

where

$$H_c = Q_c^{-1} \tag{3.6}$$

and the Lagrange multipliers (or shadow prices) associated with (3.1) are given by

$$\underline{\lambda}_c = -(N^T H_c N)^{-1} (N^T \underline{x}^d - \underline{b}). \qquad (3.7)$$

Proof

Writing the Lagrangian associated with (3.1)

$$L_c(\underline{x}, \underline{\lambda}) = q_c(\underline{x}) - \langle N^T \underline{x} - \underline{b}, \underline{\lambda} \rangle \qquad (3.8)$$

both (3.5) and (3.6) can be derived straightfordly from the following first order optimality conditions

$$Q_c(\underline{x}_c - \underline{x}^d) - N \underline{\lambda}_c = \underline{0} \qquad (3.9)$$

$$N^T \underline{x}_c - \underline{b} = \underline{0}. \qquad (3.10)$$

□

In subsequent sections the positive definiteness of Q_c and the restrictive structure of (3.4) are relaxed. However, expressions (3.5) - (3.10) are still used to study various projection properties of the method.

4. The Iterative Method for Determining the Weighting Matrix

Let $\Omega \subset E^n$ denote the set of admissible values of $\underline{x}$ from the point of view of the decision maker. Given the desired value $\underline{x}^d$ and the region R, the "nearness" of the solution of (3.1) to $\underline{x}^d$ is only affected by the weighting matrix Q_c. Q_c is initially assumed to be positive definite and subsequently relaxed to be positive semi-definite. In the latter case, the target values are allowed to be a subset of the vector $\underline{x}$. In the former case, $\underline{x}$ and the targets are identical. Different values of this matrix define different points on R as the nearest point to $\underline{x}^d$. Thus, given $\underline{x}^d$, the only way of producing a solution of (3.1) that also satisfies Ω is to re-specify Q_c. It should be noted that, in contrast to the algebraic equalities and inequalities describing R, the set Ω is assumed to exist only in the mind of the rational decision maker. It is also assumed that

$$\Omega \cap R \neq \emptyset \qquad (4.1)$$

and that this intersection is convex. Clearly, this assumption is satisfied for convex R and Ω. Thus, problem (3.1) has to be solved a number of times, by re-specifying Q_c, until a solution is found such that $\underline{x}_c \in \Omega \cap R$.

Let Q_c denote the current weighting matrix of (3.2). The solution of (3.1) using this matrix will be called the "current" optimal solution $\underline{x}_c$. This solution is presented to the decision maker, who may decide that $\underline{x}_c \notin \Omega$ since some of the elements of $\underline{x}_c$ are not quite what he wants them to be. Consequently, an alternative solution of (3.1) has to be obtained by altering Q_c. In order to do this, the policy maker is required to specify an $\underline{x}$ value which he would prefer instead of $\underline{x}_c$. This "preferred" value is denoted by $\underline{x}_p$ and obviously satisfies

$$\underline{x}_p \in \Omega \tag{4.2}$$

but <u>not necessarily</u>

$$\underline{x}_p \in R\,. \tag{4.3}$$

Thus, $\underline{x}_p$ incorporates all the alterations to $\underline{x}_c$ such that $\underline{x}_p$ is preferred, by the decision maker, to $\underline{x}_c$. Given $\underline{x}_c$ and $\underline{x}_p$, the displacement, or correction vector is defined as

$$\underline{\delta} \triangleq \underline{x}_p - \underline{x}_c\,. \tag{4.4}$$

Given $\underline{\delta}$, Q_c is altered to obtain the new weighting matrix Q_n, using the rank-one correction

$$Q_n = Q_c + \mu \frac{Q_c \underline{\delta}\, \underline{\delta}^T Q_c}{\langle \underline{\delta}, Q_c \underline{\delta} \rangle} \tag{4.5}$$

where $\mu \geq 0$ is a scalar chosen to reflect the amphasis to be given to the update. The new matrix Q_n is used in (3.1) to compute a "new" optimal solution $\underline{x}_n$. The role of μ in determining $\underline{x}_n$ and the desirable characteristics of $\underline{x}_n$, including the fact that it is an improvement on $\underline{x}_c$, are discussed in Sections 5 and 6. A single update of the form (4.5) does not necessarily yield a new optimal solution such at $\underline{x}_n \in \Omega$. Thus, (4.5) has to be utilized iteratively until $\underline{x}_n \in \Omega$. This iterative procedure, which is our attempt to formalize the Frischian scheme, is summarized below.

<u>Step 0</u> : Given the desired value $\underline{x}^d$ and the feasible region R, assume some initial symmetric, positive semi-definite weighting matrix Q_c (corresponding to (a) in Section 2).

<u>Step 1</u> : Using Q_c, solve (3.1) to obtain $\underline{x}_c$ (corresponding to (b) in Section 2).

<u>Step 2</u> : [This step describes the interaction between the policy maker and the method.] If, according to the policy maker $\underline{x}_c \varepsilon \Omega$, stop. Otherwise ask the policy maker to specify the changes, δ, required in the current optimal solution $\underline{x}_c$ to make it acceptable. The preferred value $\underline{x}_p$ is thus specified as

$$\underline{x}_p = \underline{x}_c + \underline{\delta}.$$

Alternatively, the policy maker might choose to specify $\underline{x}_p$ directly (corresponding to (c) in Section 2).

<u>Step 3</u> : Given Q_c and $\underline{\delta}$, choose a $\mu \geq 0$ and compute Q_n using (4.5) (corresponding to (c) in Section 2).

<u>Step 4</u> : Set $Q_c = Q_n$ and to Step 1 (corresponds to (f) in Section 2).

The choice of μ is bounded from above and this bound is discussed in Section 5. Also, the denominator of (4.5) has to be protected from becoming zero. This can be accomplished with small changes to $\underline{\delta}$ within limits acceptable to the policy maker. Provided this is done, Q_n remains symmetric positive (semi-)definite if Q_c is symmetric positive (semi-) definite.

5. Properties of the Method: Linear Equality Constraints

In this section the specific case arising when R is given by (3.4) is discussed. Convex and general nonlinear constraints are considered in Section 6. However, it should be noted that the basic results for linear equality constraints in this section are also essential for the discussion in Section 6. The properties of the iterative procedure of the previous section can best be analyzed by inspecting the effect of updating Q_c on the optimal solution of (3.1). The next two theorems characterize the effect of (4.5) as a corresponding update on the current optimal solution $\underline{x}_c$ and its corresponding shadow price $\underline{\lambda}_c$.

Theorem 1

Let Q_c be positive definite and the feasible region R be given by (3.4). Then for any $\underline{\delta} \neq \underline{0}$ and $\mu \geq 0$, Q_n given by (4.5) is positive definite and the new optimal solution obtained by using Q_n instead of Q_c in (3.1) is given by

$$\underline{x}_n = \underline{x}_c + \alpha P \underline{\delta} \tag{5.1}$$

where

$$P = I - H_c N(N^T H_c N)^{-1} N^T \tag{5.2}$$

$$\alpha = - \frac{\mu < \underline{\delta}, Q_c(\underline{x}_c - \underline{x}^d) >}{< \underline{\delta}, Q_c \underline{\delta} > + \mu < Q_c \underline{\delta}, P H_c (Q_c \underline{\delta}) >} \tag{5.3}$$

The corresponding shadow price vector at $\underline{x}_n$ is given by

$$\underline{\lambda}_n = \underline{\lambda}_c - \alpha (N^T H_c N)^{-1} N^T \underline{\delta} . \tag{5.4}$$

Furthermore, $\alpha \geq 0$ for $<\underline{\delta}, Q_c(\underline{x}_c - \underline{x}^d)> \leq 0$.

Proof

The fact that Q_n is positive definite if Q_c is positive definite follows trivially from (4.5). Expressions (5.1) and (3.2) can be derived, as discussed in Rustem, Velupillai, Westcott (1978, Lemmas 1 and 2), by using (4.5) instead of Q_c in (3.5) and by applying the Sherman-Morrison formula (see, e.g. Householder, 1964). Alternatively, since from (3.5) and (3.7) it follows that

$$\underline{x}_n = \underline{x}^d + H_n N \underline{\lambda}_n , \tag{5.5}$$

(5.1) and (5.2) can be derived from the derivation of $\underline{\lambda}_n$ given below. However, this will not be discussed any further. To derive the update (5.4), consider

$$\lambda_n = - (N^T H_n N)^{-1} (N^T \underline{x}^d - \underline{b}) \tag{5.6}$$

$$= - \left(N^T \left(Q_c + \mu \frac{Q_c \underline{\delta}\, \underline{\delta}^T Q_c}{(\delta^T Q_c \underline{\delta})} \right)^{-1} N \right)^{-1} (N^T \underline{x}^d - \underline{b})$$

$$= - \left(N^T H_c N - \frac{\mu}{(1+\mu)} \frac{N^T \underline{\delta}\, \underline{\delta}^T N}{<\underline{\delta}, Q_c \underline{\delta}>} \right)^{-1} (N^T \underline{x}^d - \underline{b})$$

$$= - (N^T H_c N)^{-1} (N^T \underline{x}^d - \underline{b}) - \frac{\mu (N^T H_c N)^{-1} N^T \underline{\delta}\, \underline{\delta}^T N (N^T H_c N)^{-1}}{(\mu+1) <\underline{\delta}, Q_c \underline{\delta}> - \mu <\underline{\delta}, N (N^T H_c N)^{-1} N^T \underline{\delta}>} (N^T \underline{x}^d - \underline{b})$$

$$= \underline{\lambda}_c - \frac{\mu < \underline{\delta}, N (N^T H_c N)^{-1} (N^T \underline{x}^d - \underline{b}) >}{\mu < Q_c \underline{\delta}, P H_c (Q_c \underline{\delta}) > + <\underline{\delta}, Q_c \underline{\delta}>} (N^T H_c N)^{-1} N^T \underline{\delta} . \tag{5.7}$$

Expressions (3.7) and (3.9) may be used to simplify (5.7) since

$$- < \underline{\delta}, N(N^T H_c N)^{-1} (N^T \underline{x}^d - \underline{b}) > \; = \; < \underline{\delta}, N \underline{\lambda}_c >$$

$$= \; < \underline{\delta}, Q_c (\underline{x}_c - \underline{x}^d) > \tag{5.8}$$

and (5.7) with (5.8) yields the required result (5.4) with α given by (5.3). Expression (5.7) can also be used along with (5.5) to derive (5.1). Finally, it can be seen from (5.3) by inspection, that if inequality

$$< \underline{\delta}, Q_c (\underline{x}_c - \underline{x}^d) > \; \leq 0 \tag{5.9}$$

is satisfied, then $\alpha \geq 0$, since $P H_c$ is symmetric positive semi-definite.

□

The matrix (5.2) is a well-known projection operator which projects vectors in E^n on to the subspace

$$R_o \overset{\Delta}{=} \{\underline{x} \in E^n \mid N^T \underline{x} = \underline{0}\} \tag{5.10}$$

(see, e.g. Goldfarb, 1969; Rustem, 1981). Thus, the correction term $\alpha P \underline{\delta}$ in (5.1) is along the projection of $\alpha \underline{\delta}$ on to R_o. Following the discussion in Section 3, this implies that $\underline{x}_n$ is as "near" to the preferred solution $\underline{x}_p$ as allowed by the feasible region R. The magnitude of the stepsize α can be controlled using μ. The inequality (5.9) may be interpreted as a "rationality condition" on the choice of $\underline{\delta}$. The reason for this lies in the form of (5.1). When (5.9) holds, then $\alpha \geq 0$ and thus the modification $\alpha P \underline{\delta}$ to the current optimal solution in (5.1) lies in the same direction as $\underline{\delta}$. This is the "best" alternative to $\underline{\delta}$, in the sense of the norm (3.3), allowed by the feasible region when $\underline{\delta}$ is infeasible and lies outside R. Also, since $\underline{\delta}$ is given by (4.4), it needs to be a descent direction for $q_n(\underline{x})$ at $\underline{x}_c$. Thus, the information that $\underline{x}_p$ is actually preferred to $\underline{x}_c$ is incorporated in $q_n(\underline{x})$. Using (4.5) with (5.9) establishes this result

$$
\begin{aligned}
< \underline{\delta}, \nabla q_n(\underline{x}_c) > &= < \underline{\delta}, Q_n(\underline{x}_c - \underline{x}^d) > \\
&= < \underline{\delta}, \left(Q_c + \mu \frac{Q_c \underline{\delta} \underline{\delta}^T Q_c}{<\underline{\delta}, Q_c \underline{\delta} >} \right) (\underline{x}_c - \underline{x}^d) > \\
&= (1 + \mu) < \underline{\delta}, Q_c(\underline{x}_c - \underline{x}^d) > \\
&= (1 + \mu) < \underline{\delta}, \nabla q_c(\underline{x}_c) > \qquad (5.11)
\end{aligned}
$$

Thus, $\underline{\delta}$ is a descent direction for $q_n(\underline{x}_c)$ if it is a descent direction for $q_c(\underline{x}_c)$ and condition (5.9) ensures the latter.

The properties of the method, discussed in the next section, are dependent on the form of (5.1) which is used to show that for α bounded, using μ to control its size, the distance between $\underline{x}_n$ and $\underline{x}_p$ is less than the distance between $\underline{x}_c$ and $\underline{x}_p$ (see Lemma 5 below).

The basic limitation of Theorem 1 is that Q_c, and hence Q_n, are assumed to be strictly positive definite. As discussed in Section 1, this is a rather restricted view of policy decisions since the policy maker need not attach an objective to every element of $\underline{x}$: some elements may have an objective, whereas others may be free variables in the region R. It is worth noting that when Q_c is positive semi-definite, the update (4.5) exists when $< \underline{\delta}, Q_c \underline{\delta} > > 0$.

As mentioned in Section 4, this inequality can be maintained with small changes in $\underline{\delta}$, within limits acceptable to the policy maker. Also, if the vector $\underline{\delta}$ is always specified to express corrections only in those variables which have objectives (i.e. those for which the submatrix Q is positive definite), then it can be shown that $\underline{\delta}$ lies in the range of Q_c and thereby satisfies $<\underline{\delta}, Q_c \underline{\delta}> \ > 0$. The following result establishes analogous expressions to (5.1)-(5.4) when Q_c is positive semi-definite in general but its projection on the intersection of the constraints is positive definite.

Theorem 2

Let Q_c be positive semi-definite, the feasible region be given by (3.4) and let Z denote an $n \times (n-m)$ matrix with linearly independent columns orthogonal to the $n \times m$ matrix of linearly independent constraint normals N in (3.4) with $Z^T Q_c Z$ positive definite. Then for any $\underline{\delta}$, $<\underline{\delta}, Q_c \underline{\delta}> \ > 0$, and $\mu \geq 0$, Q_n given by (4.5) is positive semi-definite and the new optimal solution is given by

$$\underline{x}_n = \underline{x}_c + \alpha_z P_z Q_c \underline{\delta} \tag{5.12}$$

where

$$P_z = Z(Z^T Q_c Z)^{-1} Z^T \tag{5.13}$$

$$\alpha_z = - \frac{\mu < \underline{\delta}, Q_c(\underline{x}_c - \underline{x}^d)>}{<\underline{\delta}, Q_c \underline{\delta}> + \mu <Q_c \underline{\delta}, P_z Q_c \underline{\delta}>} \tag{5.14}$$

The corresponding shadow price vector at $\underline{x}_h$ is given by

$$\underline{\lambda}_n = \underline{\lambda}_c - \alpha_z (N^T N)^{-1} N^T Q_c (I - P_z Q_c) \underline{\delta} \tag{5.15}$$

Furthermore, $\alpha_z \geq 0$ for

$$< \underline{\delta},\ Q_c(\underline{x}_c - \underline{x}^d) > \ \leq \ 0. \tag{5.16}$$

Proof

The positive semi-definiteness of Q_n follows directly from (4.5) for positive semi-definite Q_c. To establish (5.12), consider the optimality condition (3.9) with Q_n, given by (4.5), replacing Q_c

$$\left(Q_c + \mu \frac{Q_c\, \underline{\delta}\, \underline{\delta}^T Q_c}{< \underline{\delta},\, Q_c\, \underline{\delta} >} \right) (\underline{x}_n - \underline{x}_c + \underline{x}_c - \underline{x}^d) - N\, \underline{\lambda}_n \ = \ \underline{0} \tag{5.17}$$

Any vector $\underline{v}$ satisfying R_o given by (5.10) can be written in the form $\underline{v} = Z\underline{w}$ where $\underline{w}$ is an $(n-m)$ vector, since $Z^T N = 0$ (see, e.g. Gill and Murray, 1978). As both $\underline{x}_n$ and $\underline{x}_c$ satisfy R, $\underline{x}_n - \underline{x}_c \in R_o$ and thus

$$\underline{x}_n - \underline{x}_c \ = \ Z\, \underline{w} \tag{5.18}$$

for some $\underline{w}$. Using (5.18) and pre-multiplying (5.17) by Z^T yields

$$Z^T \left(Q_c + \mu \frac{Q_c\, \underline{\delta}\, \underline{\delta}^T Q_c}{< \underline{\delta},\, Q_c\, \underline{\delta} >} \right) (Z\, \underline{w} + \underline{x}_c - \underline{x}^d) \ = \ \underline{0}$$

$$Z\, \underline{w} \ = \ -Z \left(Z^T \left(Q_c + \mu \frac{Q_c\, \underline{\delta}\, \underline{\delta}^T Q_c}{< \underline{\delta},\, Q_c\, \underline{\delta} >} \right) Z \right)^{-1} Z^T \left(Q_c + \mu \frac{Q_c\, \underline{\delta}\, \underline{\delta}^T Q_c}{< \underline{\delta},\, Q_c\, \underline{\delta} >} \right) (\underline{x}_c - \underline{x}^d). \tag{5.19}$$

Expression (3.9) can be used to establish the equality

$$Z^T Q_c(\underline{x}_c - \underline{x}^d) \ = \ Z^T N\, \underline{\lambda}_c \ = \ \underline{0}$$

$$\underline{x}_h - \underline{x}_c = Z\,\underline{w} = -Z\left[Z^T\left(Q_c + \mu\frac{Q_c\,\underline{\delta}\,\underline{\delta}^T Q_c}{<\underline{\delta},\,Q_c\underline{\delta}>}\right)Z\right]^{-1} Z^T\left(\mu\frac{Q_c\,\underline{\delta}\,\underline{\delta}^T Q_c}{<\underline{\delta},\,Q_c\underline{\delta}>}\right)(\underline{x}_c - \underline{x}^d)$$

and the application of the Sherman-Morrison formula yields

$$\underline{x}_h = \underline{x}_c - Z^T\left[\left(Z^T Q_c Z\right)^{-1} - \frac{\mu(Z^T Q_c Z)^{-1} Z^T Q_c\underline{\delta}\,\underline{\delta}^T Q_c Z(Z^T Q_c Z)^{-1}}{<\underline{\delta}^T, Q_c\underline{\delta}> + \mu <Q_c\underline{\delta}, P_z\,Q_c\underline{\delta}>}\right] Z^T\left[\mu\frac{Q_c\underline{\delta}\,\underline{\delta}^T Q_c}{<\underline{\delta},Q_c\underline{\delta}>}\right](\underline{x}_c - \underline{x}^d)$$

from which, after some rearranging, the required result (5.12) follows.

An alternative expression to (3.7) for $\underline{\lambda}_c$ is given by the least squares solution of $\underline{\lambda}_c$ using (3.9)

$$\underline{\lambda}_c = (N^T N)^{-1} N^T Q_c(\underline{x}_c - \underline{x}^d). \tag{5.20}$$

For Q_n, this becomes

$$\underline{\lambda}_n = (N^T N)^{-1} N^T\left[Q_c + \mu\frac{Q_c\,\underline{\delta}\,\underline{\delta}^T Q_c}{<\underline{\delta},\,Q_c\,\underline{\delta}>}\right](\underline{x}_h - \underline{x}_c + \underline{x}_c - \underline{x}^d)$$

which can be simplified using (3.9) and (5.12) to yield

$$\underline{\lambda}_n = \underline{\lambda}_c + (N^T N)^{-1} N^T\left[Q_c(\underline{x}_h - \underline{x}_c) + \mu\frac{Q_c\,\underline{\delta}\,\underline{\delta}^T Q_c}{<\underline{\delta},\,Q_c\,\underline{\delta}>}(\underline{x}_h - \underline{x}_c + \underline{x}_c - \underline{x}^d)\right]$$

$$= \underline{\lambda}_c + (N^T N)^{-1} N^T Q_c\left[\alpha_z P_z\underline{\delta} + \alpha_z\mu Q_c\underline{\delta}\frac{<Q_c\underline{\delta}, P_z Q_c\,\delta>}{<\underline{\delta},\,Q_c\,\underline{\delta}>} + \mu Q_c\underline{\delta}\frac{<\underline{\delta}, Q_c(\underline{x}_c - \underline{x}^d)>}{<\underline{\delta},\,Q_c\,\underline{\delta}>}\right].$$

After some rearrangement involving (5.14), this can be expressed as (5.15). Finally, the fact that $\alpha_z \geq 0$ when (5.16) is satisfied follows from the positive semi-definiteness of $Q_c\,P_z\,Q_c$.

□

Although Theorem 2 establishes a more general result, the projection aspect of the method is equivalently characterised both in Theorem 1 and in Theorem 2 when Q_C is positive definite. The correspondence between $\underline{x}_n$ and $\underline{\lambda}_n$ computed using the expressions derived in either of these Theorems is discussed in Lemma 2 below. When analysing the method further, the positive semi-definite case will be reduced to a positive definite problem of reduced dimension, assuming that the feasible region R bounds $q_C(\underline{x})$ from below, by considering an optimization problem in which those variables in $q_C(\underline{x})$ with zero weighting have been substituted out using some of the constraints in (3.1). If $q_C(\underline{x})$ is bounded below on R, then the weighting matrix resulting from such a substitution is positive definite. This approach is adopted in more general terms in Theorem 2 where the positive definiteness of $Z^T Q_C Z$ is required instead of the positive definiteness of Q_C. The matrix $Z^T Q_C Z$ can be regarded as the projection of Q_C on to the constraints. There are a number of ways of defining the matrix Z (see Gill and Murray, 1974). However, the most suitable is the one arising from the orthogonal decomposition of N[3] and this is discussed in detail by Gill and Murray (1978) who also study the updating of Z in the presence of linear inequality constraints.

<u>Lemma 2</u>

When Q_C is positive definite, $\underline{x}_n$ and $\underline{\lambda}_n$ computed in Theorems 1 and 2 are identical. Hence

$$P\underline{\delta} = P_Z Q_C \underline{\delta} \tag{5.21}$$

$$\alpha = \alpha_Z \tag{5.22}$$

$$-(N^T H_C N)^{-1} (N^T \underline{x}^d - \underline{b}) = (N^T N)^{-1} N^T Q_C (\underline{x}_C - \underline{x}^d) \tag{5.23}$$

$$(N^T H_C N)^{-1} N^T \underline{\delta} = (N^T N)^{-1} N^T Q_C (I - P_Z Q_C) \underline{\delta} \tag{5.24}$$

Proof

To establish that $\underline{x}_n$ computed using (5.1) and (5.12) are equivalent, (5.21) and (5.22) must be shown to hold. Since $P\underline{\delta} \in R_o$, we have

$$P\underline{\delta} = Z\underline{\omega}$$

for some (n - m) vector $\underline{\omega}$. Pre-multiplying by $Z^T Q_c$ yields

$$Z^T Q_c P\underline{\delta} = Z^T Q_c (\underline{T} - H_c N(N^T H_c N)^{-1} N^T)\underline{\delta} = Z^T Q_c Z\underline{\omega}$$

$$Z^T Q_c \underline{\delta} = Z^T Q_c Z\underline{\omega}$$

since $Z^T N = 0$. Thus (5.1) and (5.12) are equivalent. It follows immediately that (5.3) and (5.14) are also equivalent and hence (5.22) holds.

To show that (5.23) holds, consider the optimality condition (3.9). The left hand side of (5.23) may be expressed as

$$-(N^T H_c N)^{-1}(N^T \underline{x}^d - \underline{b}) = (N^T H_c N)^{-1} N^T(\underline{x}_c - \underline{x}^d)$$

$$= (N^T H_c N)^{-1} N^T H_c N\underline{\lambda}_c$$

$$= \underline{\lambda}_c$$

where $\underline{\lambda}_c$ satisfies (3.9) for N full column rank. The right hand side of (5.23) can be expressed as

$$(N^T N)^{-1} N^T G(\underline{x}_c - \underline{x}^d) = (N^T N)^{-1} N^T N\underline{\lambda}_c$$

$$= \underline{\lambda}_c$$

Thus (5.23) is satisfied. Finally, using (5.21) the required result is established for (5.24)

$$(N^T N)^{-1} N^T Q_c (I - P_Z Q_c)\underline{\delta} = (N^T N)^{-1} N^T Q_c H_c N(N^T H N)^{-1} N^T\underline{\delta}$$

$$= (N^T H_c N)^{-1} N^T \underline{\delta}$$

□

6. Properties of the Method: General Convex and Nonlinear constraints

A desirable local property of the method is that every time Q_c is updated and a new optimal solution is computed, the policy maker is more satisfied with the new optimal solution, $\underline{x}_n$, than he was with the current optimal solution, $\underline{x}_c$. The following two Theorems establish such results for the linear equality constrained case. These will then be extended to general convex and also nonlinear constraints.

Theorem 3

For $0 \le \alpha \le 2$ in (5.1), the inequality

$$\| \underline{x}_n - \underline{x}_p \|_{Q_c} \le \| \underline{x}_c - \underline{x}_p \|_{Q_c} \tag{6.1}$$

holds. Furthermore, for $0 < \alpha < 2$, (6.1) is a strict inequality.

Proof

Consider (5.1) with $\alpha = 1$

$$\underline{x}_1 = \underline{x}_c + P\,\underline{\delta} \tag{6.2}$$

Hence, $\underline{x}_1 - \underline{x}_c$ is the projection of $\underline{\delta}$ on to $\{\underline{x} \mid N^T \underline{x} = \underline{0}\}$, or $\underline{x}_1$ is the projection of $\underline{x}_p$ on to R given by (3.4). The inequality

$$< \underline{x}_p - \underline{x}_1,\ Q_c(\underline{x}_c - \underline{x}_1) > \ \le 0 \tag{6.3}$$

follows from the projection property of $\underline{x}_1$ (see, e.g. Luenberger, 1969, p.69; Rustem, 1981; Lemma 4.2). Defining $\underline{x}(\alpha)$ as

$$\underline{x}(\alpha) \triangleq \underline{x}_c + (\underline{x}_1 - \underline{x}_c) \tag{6.4}$$

for $0 \le \alpha \le 2$ and using (6.3) we have

$$\|\underline{x}(\alpha) - \underline{x}_p\|^2_{Q_c} = \|\underline{x}_c - \underline{x}_p + \alpha(\underline{x}_1 - \underline{x}_p)\|^2_{Q_c}$$

$$= \|\underline{x}_c - \underline{x}_p\|^2_{Q_c}$$

$$+ 2\alpha < \underline{x}_c - \underline{x}_1 + \underline{x}_1 - \underline{x}_p , \; Q_c(\underline{x}_1 - \underline{x}_c) >$$

$$+ \alpha^2 < \underline{x}_1 - \underline{x}_c , \; Q_c(\underline{x}_1 - \underline{x}_c)$$

$$\le \|\underline{x}_c - \underline{x}_p\|^2_{Q_c}$$

from which the result follows with $\underline{x}_n = \underline{x}(\alpha)$. Furthermore, it should be noted that for $0 < \alpha < 2$ (6.2) becomes a strict inequality.

□

Lemma 3

For $\underline{x}_1$ given by (6.3), the inequality

$$< \underline{x}_p - \underline{x}_1 , \; Q_n(\underline{x}_1 - \underline{x}_c) > \; \ge \; 0 \tag{6.5}$$

holds.

Proof

Using the update formula (4.5) in (6.8) we have

$$< \underline{x}_p - \underline{x}_1 , \; Q_n(\underline{x}_1 - \underline{x}_c) > \; = \; < \underline{x}_p - \underline{x}_1 , \; Q_c(\underline{x}_1 - \underline{x}_c) >$$

$$+ \mu \frac{\langle \underline{x}_p - \underline{x}_1, Q_c(\underline{x}_p - \underline{x}_c) \rangle \langle \underline{x}_1 - \underline{x}_c, Q_c(\underline{x}_p - \underline{x}_c) \rangle}{\langle \underline{x}_p - \underline{x}_c, Q_c(\underline{x}_p - \underline{x}_c) \rangle}$$

$$= \langle \underline{x}_p - \underline{x}_1, Q_c(\underline{x}_1 - \underline{x}_c) \rangle$$

$$+ \mu \frac{\langle \underline{x}_p - \underline{x}_1, Q_c(\underline{x}_p - \underline{x}_1 + \underline{x}_1 - \underline{x}_c) \rangle \langle \underline{x}_1 - \underline{x}_c, Q_c(\underline{x}_p - \underline{x}_1 + \underline{x}_1 - \underline{x}_c) \rangle}{\langle \underline{x}_p - \underline{x}_c, Q_c(\underline{x}_p - \underline{x}_c) \rangle}$$

$$\geq 0$$

where the last inequality follows from the projection property.

□

<u>Theorem 4</u>

For $0 \leq \alpha \leq 2$ in (5.2) the inequality

$$\| \underline{x}_p - \underline{x}_p \|_{Q_n} \leq \| \underline{x}_c - \underline{x}_p \|_{Q_n} \tag{6.6}$$

holds. Furthermore, for $0 < \alpha < 2$, (6.6) is a strict inequality.

<u>Proof</u>

Let $\underline{x}(\alpha)$ be defined by (6.4). For $0 \leq \alpha \leq 2$ and using (6.5) we have

$$\begin{aligned} \| \underline{x}_p - \underline{x}(\alpha) \|^2_{Q_n} &= \| \underline{x}_p - \underline{x}_c - \alpha(\underline{x}_1 - \underline{x}_c) \|^2_{Q_n} \\ &= \| \underline{x}_p - \underline{x}_c \|^2_{Q_n} \\ &\quad - 2\alpha \langle \underline{x}_p - \underline{x}_1 + \underline{x}_1 - \underline{x}_c, Q_n(\underline{x}_1 - \underline{x}_c) \rangle \\ &\quad + \alpha^2 \langle \underline{x}_1 - \underline{x}_c, Q_n(\underline{x}_1 - \underline{x}_c) \rangle \end{aligned}$$

$$\leq \| \underline{x}_p - \underline{x}_c \|^2_{Q_n}$$

which establishes the required result with $\underline{x}_n = \underline{x}(\alpha)$.

□

The difficulty in establishing similar results in the case of general convex constraints arises from the fact that there may not be a point between $\underline{x}_c$ and $\underline{x}_n$, along $\underline{x}_n - \underline{x}_c$, which can be expressed as the projection of $\underline{x}^d$ onto R, with respect to the weighting Q_n, and hence solves the problem.

$$\min\{ \tfrac{1}{2} < \underline{x} - \underline{x}^d, Q_n(\underline{x} - \underline{x}^d) > \mid \underline{x} \in R \} \tag{6.7}$$

with some $\mu \geq 0$ defining Q_n via (4.5) and with convex R. The importance of this becomes clear when the above linear equality constrained case is considered. If, for given μ, $\underline{x}_n$ is such that $\alpha > 2$ in (5.2), in the linear equality case, reducing μ clearly reduces α given by (5.3) and for every value of μ, $\underline{x}_n$ can be expressed in terms of (5.1). Hence, if $\alpha > 2$, μ can be reduced to define an $\underline{x}_n$ for which the bound $0 \leq \alpha \leq 2$ is satisfied.

The main concept necessary for extending the results of Theorems 3 and 4 to general convex R is the line passing through $\underline{x}_n$ and $\underline{x}_c$. By considering the projection of $\underline{x}_p$ on this line, it is shown in Theorem 5 below that (6.1) and (6.6) hold for $\underline{x}_n$ close enough to $\underline{x}_c$. It is also shown that if this is not the case, reducing μ brings $\underline{x}_n$ close to $\underline{x}_c$ so that for small enough μ (6.1) and (6.6) hold.

The following two Lemmas establish the basic results used in Theorem 5.

Lemma 4

For $\mu \geq 0$ in (4.5) and for $\underline{x}_c$ and $\underline{x}_n$ given by the solutions of (3.1) and (6.7) respectively, with convex R, the inequality

$$< \underline{x}_p - \underline{x}_c, Q_c(\underline{x}_n - \underline{x}_c) > \; \geq 0 \tag{6.8}$$

holds if $< \underline{\delta}, Q_c(\underline{x}_c - \underline{x}^d) > \; \leq 0$.

Proof

The inequality

$$< \underline{x}_n - \underline{x}_c, Q_n(\underline{x}_n - \underline{x}^d) > \; \leq 0 \tag{6.9}$$

follows from the optimality of $\underline{x}_n$ for (6.7). Using (4.5) we have

$$0 \geq < \underline{x}_n - \underline{x}_c, Q_n(\underline{x}_n - \underline{x}^d) > \; = < \underline{x}_n - \underline{x}_c, Q_c(\underline{x}_n - \underline{x}^d) > +$$

$$+ \frac{\mu}{< \underline{\delta}, Q_c\underline{\delta} >} < \underline{x}_n - \underline{x}_c, Q_c\underline{\delta} >< \underline{\delta}, Q_c(\underline{x}_n - \underline{x}^d) >$$

$$= < \underline{x}_n - \underline{x}_c, Q_c(\underline{x}_n - \underline{x}_c + \underline{x}_c - \underline{x}^d) >$$

$$+ \frac{\mu}{< \underline{\delta}, Q_c\underline{\delta} >} [< \underline{x}_n - \underline{x}_c, Q_c(\underline{x}_p - \underline{x}_c) >^2$$

$$+ < \underline{x}_c - \underline{x}^d, Q_c(\underline{x}_p - \underline{x}_c) >< \underline{x}_p - \underline{x}_c, Q_c(\underline{x}_n - \underline{x}_c) >]. \tag{6.10}$$

Since

$$< \underline{x}_n - \underline{x}_c, Q_c(\underline{x}_c - \underline{x}^d) > \; \geq 0 \tag{6.11}$$

follows from the optimality of $\underline{x}_c$ for (3.1) and

$$< \underline{x}_p - \underline{x}_c, Q_c(\underline{x}_c - \underline{x}^d) > \leq 0$$

follows by hypothesis, inequality (6.10) can only be satisfied if (6.8) holds.

□

It should be noted that condition (5.9) also plays an important role in establishing (6.8).

Lemma 5

If the conditions of Lemma 4 hold, then $\mu \to 0$ implies that $\| \underline{x}_c - \underline{x}_n \| \to 0$. Furthermore, for a fixed value of $\alpha \geq 0$, there exists a $\mu \geq 0$ that satisfies (5.3) and (5.4)

Proof

Using (4.5) with inequality (6.9) yields

$$0 \leq < \underline{x}_c - \underline{x}_n, Q_c(\underline{x}_n - \underline{x}^d) > + \frac{\mu}{< \underline{\delta}, Q_c \underline{\delta} >} < \underline{x}_c - \underline{x}_n, Q_c \underline{\delta} > < \underline{\delta}, Q_c(\underline{x}_n - \underline{x}^d) >. \quad (6.12)$$

Using (6.8) and (6.9) it can be concluded from (6.12) that for $\mu > 0$, $< \underline{\delta}, Q_c(\underline{x}_n - \underline{x}^d) > \leq 0$. Thus

$$< \underline{x}_c - \underline{x}_n, Q_c \underline{\delta} > < \underline{\delta}, Q_c(\underline{x}_n - \underline{x}^d) > \geq 0 \quad (6.13)$$

and moreover we have

$$< \underline{x}_c - \underline{x}_n, Q_c(\underline{x}_n - \underline{x}^d) > \leq 0. \quad (6.14)$$

As $\mu \to 0$, the second term in (6.12) also approaches to zero implying, through inequalities (6.12) and (6.14), that $< \underline{x}_c - \underline{x}_n, Q_c(\underline{x}_n - \underline{x}^d) > \to 0$. Since $\| \underline{x}_n - \underline{x}^d \| \neq 0$, we have $\| \underline{x}_c - \underline{x}_n \| \to 0$. As $\underline{x}_n \neq \underline{x}^d$,

$<\underline{x}_c - \underline{x}_n, Q_c(\underline{x}_n - \underline{x}^d)> = 0$ due to orthogonality implies that $\underline{x}_n$ solves (3.1) and hence $\underline{x}_n = \underline{x}_c$. Finally, consider α given by (5.3) or (5.4). Due to the equivalence of both expressions (see Lemma 2), only (5.3) is considered. Given α, $\mu = -\alpha < \underline{\delta}, Q_c \underline{\delta} > / (\alpha < \underline{\delta}, Q_c P \underline{\delta}> + < \underline{\delta}, Q_c(\underline{x}_c - \underline{x}^d) >)$. Thus, we only have to show that the denominator of this expression is negative. Using (5.1), the denominator can be written as $<\underline{\delta}, Q_c(\underline{x}_n - \underline{x}_c + \underline{x}_c - \underline{x}^d)> = <\underline{\delta}, Q_c(\underline{x}_n - \underline{x}^d)> \leq 0$ It also follows from (6.12) - (6.14) that for a strict inequality in (5.9) the above inequality is strict. Hence the denominator of the expression for μ is strictly negative. Thus, for a fixed $\alpha \geq 0$ the required result follows.

□

Lemma 5 implies that the scalar μ can be used by the policy maker to control the size of the norm $\|\underline{x}_c - \underline{x}_n\|$. This result is utilised in Theorem 5.

Theorem 5

Let the conditions of Lemma 4 be satisified and consider the line passing through $\underline{x}_c$ and $\underline{x}_n$. Let $\underline{x}_1$ be the projection of $\underline{x}_p$ on to this line. Thus

$$\underline{x}_1 = \underline{x}_c + P\underline{\delta}$$

where P is the operator projecting, under the norm $\|\cdot\|_{Q_c}$, vectors in E^n on to this line. Then for $\underline{x}(\alpha)$ given by
$\underline{x}(\alpha) \triangleq \underline{x}_c + \alpha(\underline{x}_1 - \underline{x}_c)$ and $0 \leq \alpha \leq 2$

$$\|\underline{x}(\alpha) - \underline{x}_p\|_{Q_c} \leq \|\underline{x}_c - \underline{x}_p\|_{Q_c} \tag{6.15}$$

$$\|\underline{x}(\alpha) - \underline{x}_p\|_{Q_n} \leq \|\underline{x}_c - \underline{x}_p\|_{Q_n} \tag{6.16}$$

Also

$$\underline{x}_n - \underline{x}_c = \tau(\underline{x}_1 - \underline{x}_c) \tag{6.17}$$

with $\tau \geq 0$, which implies that both $\underline{x}_1 - \underline{x}_c$ and $\underline{x}_n - \underline{x}_c$ lie in the same direction. Furthermore, if

$\underline{x}_n \notin X \triangleq \{ \underline{x} \mid \underline{x} = \underline{x}_c + \alpha(\underline{x}_1 - \underline{x}_c),\ \alpha \in [0, 2] \}$ then there exists a $\mu \geq 0$, small enough, so that the resulting solution of (6.7) satisfies X and thereby (6.15) and (6.16) with $\underline{x}_n = \underline{x}(\alpha)$.

□

Proof

Inequalities (6.15) and (6.16) are established in exactly the same way as in Theorems 3 and 4, with the feasible region R replaced by the line passing through $\underline{x}_c$ and $\underline{x}_n$.

To show that $\tau \geq 0$ in (6.17), consider (6.8) which yields

$$0 \leq < \underline{x}_p - \underline{x}_c,\ Q_c(\underline{x}_n - \underline{x}_c) > = \tau < \underline{x}_p - \underline{x}_c,\ Q_c(\underline{x}_1 - \underline{x}_c) >. \tag{6.18}$$

Since the inequality

$$< \underline{x}_p - \underline{x}_1 + \underline{x}_1 - \underline{x}_c,\ \ Q_c(x_1 - x_c) > \geq 0$$

follows from (6.3), in order to preserve the non-negativity of (6.18), the condition $\tau \geq 0$ must be satisfied. Thus, if $\underline{x}_n$ can be expressed as

$$\underline{x}_n = \underline{x}(\alpha) = \underline{x}_c + \alpha(\underline{x}_1 - \underline{x}_c) \tag{6.18}$$

for some $\alpha \in [0,2]$ then (6.15) and (6.6) hold. However, if $\underline{x}_n$ is further away from $\underline{x}_c$ so that $\alpha > 2$, then μ can be reduced. According to Lemma 5, $\mu \to 0$ implies $\| \underline{x}_n - \underline{x}_c \| \to 0$. In view of (6.18), $\| \underline{x}_n - \underline{x}_c \| \to 0$ implies that $\alpha \to 0$ since $\| \underline{x}_1 - \underline{x}_c \|$ is not related to μ.

□

The extension of these results to general nonlinear constraints is possible since reducing μ brings Q_n closer to Q_c, independently of Lemmas 4 and 5. Thus, there exists a $\mu > 0$ for which inequalities (6.9) and (6.11) are satisfied, and thence the results of Theorem 5 hold for general nonlinear constraints.

The above results are particularly helpful if $\underline{x}_c$ is at a vertex of linear constraints. If such vertices can be excluded, the following Theorem provides an *alternative* characterisation of the results of Theorem 5 in the presence of general nonlinear constraints.

<u>Theorem 6</u>

There exist scalars $\mu \geq 0$ ($\mu > 0$ with vertices excluded) and $\delta(\mu) \geq 0$ ($\delta(\mu) > 0$ excluding vertices) such that $\| \underline{x}_n - \underline{x}_c \| \leq \delta(\mu)$ and all inequality constraints satisfied as equalities (i.e. active) at $\underline{x}_c$ are the same constraints as those satisfied as equalities at $\underline{x}_n$. Using the mean value theorem, these active constraints may be expressed as $\underline{g}(\underline{x}_c) - \underline{g}(\underline{x}_n) = \underline{0} = N^T(\underline{x}_c, \underline{x}_n)(\underline{x}_c - \underline{x}_n)$ where $\underline{g}$ is the vector of m_a active constraints and

$N(\underline{x}_c, \underline{x}_n) \triangleq [\nabla g_1(\underline{x}_c - t_1(\underline{x}_n - \underline{x}_c)), \ldots, \nabla g_{m_a}(\underline{x}_c - t_{m_a}(\underline{x}_n - \underline{x}_c))]$

with $t_1, \ldots, t_{m_a} \in (0,1)$. The line connecting $\underline{x}_c$ and $\underline{x}_n$ clearly satisfies the above linear equality and thus, for small μ, the results of Theorem 5 are valid for general nonlinear inequality constraints.

<u>Proof</u>

The linear equalities for $\underline{g}(\underline{x})$ are obtained by a simple application of the mean value theorem (see Ortega and Rheinboldt, 1970; Theorem 3.2.2). In order to show that for small μ, the active constraints at

$\underline{x}_c$ are the same as those active at $\underline{x}_n$ consider the first order optimality conditions for (3.1) and (6.7) when R is given by

$$\{ \underline{x} \in E^n \mid \bar{\underline{g}}(\underline{x}) \leq \underline{0} \}.$$

with $\bar{\underline{g}}$ as the vector of nonlinear inequality constraints. For (3.1) these conditions are

$$Q_c(\underline{x}_c - \underline{x}^d) + N_c \underline{\lambda}_c = \underline{0}$$

$$\bar{\underline{g}}(\underline{x}_c) \leq \underline{0}, \quad \langle \bar{\underline{g}}(\underline{x}_c), \underline{\lambda}_c \rangle = 0, \quad \underline{\lambda}_c \geq \underline{0}$$

where N_c is the matrix of constraint normals evaluated at $\underline{x}_c$ and for (6.7) we have

$$\left(Q_c + \mu \frac{Q_c \underline{\delta}\, \underline{\delta}^T Q_c}{\underline{\delta}^T Q_c \underline{\delta}} \right) (\underline{x}_n - \underline{x}^d) + N_n \underline{\lambda}_n = \underline{0}$$

$$\bar{\underline{g}}(\underline{x}_n) \leq \underline{0}, \quad \langle \bar{\underline{g}}(\underline{x}_n), \quad \underline{\lambda}_n \rangle = 0, \quad \underline{\lambda}_n \geq \underline{0}$$

Thus, as $\mu \to 0$, $\underline{x}_n \to \underline{x}_c$ and with strict complementarity holding at both $\underline{x}_c$ and $\underline{x}_n$, the active constraints at $\underline{x}_c$ predict the active constraints at $\underline{x}_n$ and vice versa for some $\mu > 0$, since $\lambda_n^i > 0 \Rightarrow \lambda_c^i > 0$ for the active constraint i as $\underline{x}_n \to \underline{x}_c$.

Having constructed the intersection of hyperplanes which characterise the active constraints between $\underline{x}_n$ and $\underline{x}_c$ and in which the line through $\underline{x}_n$ and $\underline{x}_c$ passes, the rest of the proof is identical to Theorem 5.

□

7. Khatchian's Ellipsoid Algorithm and the Complexity of the Policy Design Process

In this section we discuss briefly the termination property of a slightly modified and less intuitive version of the policy design process outlined in previous sections. In particular, we show that the modified policy design process terminates after a finite number of iterations if the econometric model and the inequality constraints bounding the region Ω are assumed to be linear. The region Ω and the inequality constraints bounding it are still assumed to exist only in the mind of the policy maker. The consideration of the convergence of policy design processes under such circumstances may, in reality, be a contradiction in itself. Indeed, as discussed below, the convergence of the method ultimately depends on the policy designer and he/she may arbitrarily extend or truncate this process. Nevertheless, the "condition" under which the method converges, does provide an insight to the method from a different vantage point and indicates the reason why the policy designer might extend or truncate the process. This is done by establishing an equivalence of the algorithms in Section 4 with Khatchian's (1979, 1980) ellipsoid algorithm for linear programming. The latter algorithm has been shown to terminate in polynomial time (i.e. the number of iterations required to arrive at a solution - or to establish the absence of one - is bounded by a polynomial in the original data of the problem (Khatchian, 1979, 1980; Kozlov, Tarasov and Khatchian, 1980; Aspvall and Stone, 1979). This result is summarised in Theorem 7 below.

In order to introduce the ellipsoid algorithm, consider first the problem of finding a feasible point satisfying the following system of inequalities

$$\langle \underline{h}_i, \underline{x} \rangle \leq g_i \qquad i = 1, \ldots, p \tag{7.1}$$

where $\underline{h}_i \in E^n$, $p \geq 2$, $n \geq 2$. Khatchian's algorithm, summarised below, finds such a point, or establishes its nonexistence, in a finite number of iterations. Let L be the length of the binary encoding of the input data $\underline{h}_i$, g_i, $i = 1, \ldots, p$, i.e. the number of 0's and 1's needed to write these coefficients in binary form:

$$L = \sum_{i,j=1}^{n,p} \log_2(|h_{ij}|+1) + \sum_{j=1}^{p} \log_2(|g_j|+1) + \log_2 np + 2. \tag{7.2}$$

where h_{ij} is the j th element of vector $\underline{h}_i$. Khatchian's algorithm assumes that coefficients h_{ij}, g_j are integers. This can trivially be achieved, in general, by suitably scaling each inequality. The algorithm discussed below can also be used directly for non-integer h_{ij}, g_j. In this case, the slight change in the properties of the algorithm are discussed in Goldfarb and Todd (1982).

<u>Khatchian's Algorithm</u>

<u>Step 1</u> : (Initialisation) Set $\underline{x}_o = 1$, $H_o = 2^{2L} I$, $k = 0$

<u>Step 2</u> : If $\underline{x}_k$ satisfies

$$\langle \underline{h}_i, \underline{x}_k \rangle < g_i + 2^{-L} \qquad \forall\, i = 1, \ldots p \tag{7.3}$$

then terminate the algorithm with $\underline{x}_k$ as a feasible solution. If $k < 4(n+1)^2 L$, then go to Step 3. Otherwise, terminate the algorithm responding that no solution exists.

Step 3 : Select any inequality for which

$$< \underline{h}_i, \underline{x}_k > \; \geq \; g_i + 2^{-L} \qquad (7.4)$$

and set

$$\underline{x}_{k+1} = \underline{x}_k - \frac{1}{(n+1)} \frac{H_k \underline{h}_i}{< \underline{h}_i, H_k \underline{h}_i >^{\frac{1}{2}}} \qquad (7.5)$$

and

$$H_{k+1} = \frac{n^2}{n^2 - 1} \left[H_k - \frac{2}{n+1} \frac{H_k \underline{h}_i \underline{h}_i^T H_k}{< \underline{h}_i, H_k \underline{h}_i >} \right] \qquad (7.6)$$

Set $k = k + 1$ and go to Step 2.

It can easily be shown that H_{k+1} is symmetric positive definite if H_k has these properties (see Aspvall and Stone, 1979; Lemma 3). Thus, there is no danger of the denominator of the above expressions to vanish, provided $\underline{h}_i \neq \underline{0}$. The algorithm above actually finds a feasible solution for the system of inequalities (7.3). However, the following Lemma ensures that this is compatible with the requirement of the system (7.1).

Lemma 6 (Aspvall and Stone, 1979; Lemma 6)

The system of inequalities (7.1) has a solution if, and only if, the system of strict inequalities (7.3) has a solution.

□

The following theorem implies that the above algorithm returns a feasible solution or establishes the non-existence of one in at most $4(n+1)^2 L$ iterations.

Theorem 7 (Aspvall and Stone, 1979; Theorem 1)

The above algorithm returns a feasible solution if and only if (5.1) is satisfiable.

□

Consider now the problem of finding the solution to the system of inequalities (7.1) in the presence of linear equalities

$$N^T \underline{x} = \underline{b} . \tag{7.7}$$

In this case, given a starting point $\tilde{\underline{x}}_o$, the initial solution estimate is defined by

$$\underline{x}_o = \tilde{\underline{x}}_o - H_k N(N^T H_k N)^{-1} (N^T \tilde{\underline{x}}_o - \underline{b})$$

It can be verified that $N^T \underline{x}_o = \underline{b}$. In order that all $\underline{x}_k$ generated by the algorithms satisfy $N^T \underline{x}_k = \underline{b}$, we have to replace (5.5) by

$$\underline{x}_{k+1} = \underline{x}_k - \frac{1}{n+1} \frac{P_k H_k \underline{h}_i}{\langle \underline{h}_i , H_k \underline{h}_i \rangle^{\frac{1}{2}}} \tag{7.8}$$

where $P_k = I - H_k N(N^T H_k N)^{-1} N^T$, similar to (5.2), or, with $H_k = Q_k^{-1}$,

$P_k = Z(Z^T Q_k Z)^{-1} Z_k^T Q_k$, similar to (5.12)-(5.13)[13]. It can also be verified that $P_k(\underline{x}_{k+1} - \underline{x}_k) = \underline{x}_{k+1} - \underline{x}_k$. All the other steps of the algorithm remain unchanged. Thus, any feasible point generated by the algorithm also satisfies the linear equalities given by (7.7).

Let us now return to the Algorithm in Section 4 and attempt to identify the reason why the policy maker may wish to specify a given direction $\underline{\delta}$ as a direction of "improvement". When the policy maker is asked in Step 2 of the algorithm in Section 4 to specify $\underline{\delta}$ or the preferred value $\underline{x}_p$ $(= \underline{x}_c + \underline{\delta})$, he is, in effect, required to specify the point, nearest to $\underline{x}_c$, which is in Ω. This "nearness" is measured with respect to the current weighting matrix Q_c. Thus, he is asked to specify $\underline{x}_p$, which is the solution of

$$\min\{ \tfrac{1}{2} \| \underline{x} - \underline{x}_c \|^2_{Q_c} \quad | \quad \langle \underline{h}, \underline{x} \rangle \leq g \} \tag{7.9}$$

where $\langle \underline{h}, \underline{x} \rangle \leq g$ is one of the implicit constraints describing the region Ω and violated at $\underline{x}_c$. The policy maker may not know that such a constraint exists until he notices that x_c is violating it (i.e. $\langle \underline{h}, \underline{x}_c \rangle > g$).

Clearly, the policy maker also does not know $\underline{h}$ and g but can only specify $\underline{x}_p$. It is argued below that this is sufficient to identify $\underline{h}$ to some degree, and thereby quantify Ω, if $\underline{x}_p$ is interpreted as the solution of (7.9). This interpretation is shown to allow the use of Khatchian's algorithm, discussed above, to solve for a feasible point of $\Omega \cap R$ in polynomial time by updating $H = Q^{-1}$ and using (7.5)-(7.6).

The solution of (7.9) can be obtained by writing the first order necessary conditions of optimality

$$\begin{aligned} \underline{x}_p - \underline{x}_c &= - Q_c^{-1} \underline{h} \lambda \\ \langle \underline{h}, \underline{x}_p \rangle \leq g, \quad \lambda(\langle \underline{h}, \underline{x}_p \rangle - g) &= 0 \\ \lambda &\geq 0 . \end{aligned} \tag{7.10}$$

Thus $\underline{\delta} = \underline{x}_p - \underline{x}_c = - Q_c^{-1} \underline{h} \lambda$ where the Lagrange multiplier (shadow price) λ is non-negative. It can be seen from (7.10) that

$$\lambda = - < \underline{h}, \underline{x}_p - \underline{x}_c > / < \underline{h}, Q_c^{-1} \underline{h} >$$

or

$$\lambda = - < \underline{x}_p - \underline{x}_c, Q_c(\underline{x}_p - \underline{x}_c) > / < \underline{h}, \underline{x}_p - \underline{x}_c > .$$

Since $\lambda \geq 0$, these yield

$$\lambda = \frac{< \underline{x}_p - \underline{x}_c, Q_c(\underline{x}_p - \underline{x}_c) >^{\frac{1}{2}}}{< \underline{h}, Q_c^{-1} \underline{h} >^{\frac{1}{2}}}$$

$$\triangleq \frac{\| \underline{\delta} \|_{Q_c}}{\| \underline{h} \|_{Q_c^{-1}}}$$

Thus, we have

$$\frac{\underline{\delta}}{\| \underline{\delta} \|_{Q_c}} = - \frac{Q_c^{-1} \underline{h}}{\| \underline{h} \|_{Q_c^{-1}}} . \qquad (7.11)$$

It may well be that more than one constraint is violated at $\underline{x}_c$. Assume that the system

$$H^T \underline{x} \leq \underline{g}$$

is violated at $\underline{x}_c$ (i.e. $H^T \underline{x}_c > \underline{g}$) for some appropriate dimensional matrix H and vector $\underline{g}$. Then (7.9) can be rewritten as

$$\min \{ \tfrac{1}{2} \| \underline{x} - \underline{x}_c \|^2_{Q_c} \mid H^T \underline{x} \leq \underline{g} \} \qquad (7.12)$$

for which the first order optimality conditions are

$$\underline{x}_p - \underline{x}_c = -Q_c^{-1} H \underline{\lambda} \tag{7.13}$$

$$< \underline{\lambda}, H^T \underline{x}_p - \underline{g}) > = 0, \quad H^T \underline{x}_p \leq \underline{g}$$

$$\underline{\lambda} \geq \underline{0}$$

for an appropriate dimensional Lagrange multiplier $\underline{\lambda}$. Thus, $\underline{\delta} = - Q_c^{-1} H \underline{\lambda}$ and $\underline{\lambda}$ is given by the solution of the quadratic programming problem (7.12). Khatchian's algorithm can utilize the constraints $H^T \underline{x} \leq \underline{g}$ using $\underline{\lambda}$ by defining $\underline{h} \triangleq H \underline{\lambda}$, $g = < \underline{g}, \underline{\lambda} >$. This is called a "surrogate" cut and is discussed in Goldfarb and Todd (1982) and Bland, Goldfarb and Todd (1981).

The above discussion illustrates the correspondence of the correction vector $\underline{\delta}$ used in the Algorithm in Section 4 and the constraint normals $\underline{h}$ used in Khatchian's algorithm. We now reformulate the the rank-one update (4.5) used in Section 4 as

$$Q_n = \frac{n^2 - 1}{n^2} \left[Q_c + \frac{2}{n-1} \frac{Q_c \underline{\delta} \underline{\delta}^T Q_c}{< \underline{\delta}, Q_c \underline{\delta} >} \right] \tag{7.14}$$

where μ in (4.5) is given by $\mu = 2/(n-1)$ and $n \geq 2$. In this case, the results of Theorem 1 can be formulated such that (5.1) remains unchanged and α is given by

$$\alpha = \frac{2 < \underline{\delta}, Q_c (\underline{x}_c - \underline{x}^d) >}{(n-1) < \underline{\delta}, Q_c \underline{\delta} > + 2 < Q_c \underline{\delta}, PH_c (Q_c \underline{\delta}) >} \tag{7.15}$$

The value $\mu = 2/(n-1)$ eliminates the need for specifying μ explicitly. The vector of shadow prices $\underline{\lambda}_n$ (5.4) is replaced by

$$\underline{\lambda}_n = \frac{n^2 - 1}{n} \left[\underline{\lambda}_c - \alpha (N^T H_c N)^{-1} N^T \underline{\delta} \right].$$

Similar results can also be derived for (5.12) - (5.15) of Theorem 2.

The inverse of Q_n can be written using (7.14) to be

$$H_n = \frac{n^2}{n^2 - 1} \left[H_c - \frac{2}{(n+1)} \frac{\underline{\delta}\,\underline{\delta}^T}{<\underline{\delta}, Q_c \underline{\delta}>} \right] \tag{7.16}$$

Using the equivalence between $\underline{\delta}$ and $\underline{h}$ given by (7.11), we can express (5.1) as

$$\underline{x}_n = \underline{x}_c - \alpha P H_c \frac{\underline{h}}{<\underline{h}, H_c \underline{h}>^{\frac{1}{2}}} \tag{7.17}$$

where

$$\alpha = - \frac{\mu <\underline{h}, \underline{H}_c \underline{h}>^{\frac{1}{2}} <\underline{h}, (\underline{x}_c - \underline{x}^d)>}{<\underline{h}, H_c \underline{h}> + \mu <\underline{h}, P H_c \underline{h}>} \tag{7.18}$$

and (7.16) becomes

$$H_n = \frac{n^2}{n^2 - 1} [H_c - \frac{2}{(n+1)} \frac{H_c \underline{h}\,\underline{h}^T H_c}{<\underline{h}, H_c \underline{h}>}] \tag{7.19}$$

we note that expressions (7.6) and (7.8) of Khatchian's algorithm are identical to (7.19) and (7.17) respectively, with $P_k = P$, $H_k = H_c$, $H_{k+1} = H_n$, $\underline{x}_{k+1} = \underline{x}_n$, $\underline{x}_k = \underline{x}_c$, and α set to

$$\alpha = 1/(n+1)$$

Also, in the case when $\underline{\delta}$ is given by (7.13), we have

$$\underline{x}_n = \underline{x}_c - \alpha \frac{P H_c \mathcal{H} \underline{\lambda}}{<\mathcal{H}\underline{\lambda}, H_c \mathcal{H}\underline{\lambda}>^{\frac{1}{2}}}$$

replacing (7.17). Similarly, (7.8) is replaced by

$$\underline{x}_{k+1} = \underline{x}_k - \frac{1}{n+1} \frac{P_k H_k \mathcal{H} \underline{\lambda}}{< \mathcal{H} \underline{\lambda}, H_k \mathcal{H} \underline{\lambda} >^{\frac{1}{2}}} .$$

Thus, by exploring the similarities between Khatchian's algorithm in the presence of equality constraints and the algorithm in Section 4, we have ended up with an algorithm which has specific values for μ and α that, by invoking Theorem 7, guarantee termination in a finite number of steps or iterations. Assuming that $R \cap \Omega \neq \emptyset$, we summarise the algorithm:

<u>Step 0</u> : Given $\underline{x}_o$ and the equality constraints, set $H_o = 2^{2L} I$, $k = 0$ and compute $\underline{x}_c = \tilde{\underline{x}}_o - H_o N(N^T H_o N)^{-1} (N^T \tilde{\underline{x}}_o - \underline{b})$.

<u>Step 1</u> : If $\underline{x}_c \in \Omega$ <u>stop</u>. Otherwise ask the policy maker to specify $\underline{\delta}$.

<u>Step 2</u> : Compute

$$\underline{x}_n = \underline{x}_c + \frac{1}{n+1} \frac{P \underline{\delta}}{< \underline{\delta}, Q_c \underline{\delta} >}$$

and H_n using (5.16). Set $\underline{x}_c = \underline{x}_n$, $H_c = H_n$ and go to step 1.

In the above algorithm, $\underline{x}$ is not the solution to the optimization problem (3.1). However, $\underline{x}_n$ still exhibits the same property as in (5.1) as $\underline{x}_n - \underline{x}_c$ lies along the projection of the direction $\underline{\delta}$ specified by the policy maker. In addition, the above algorithm ensures finite termination in polynomial time by invoking Theorem 7.

The above algorithm and the associated concept of finite termination provides an insight to the policy decision process. Nevertheless, this finite termination property may easily be undermined by an indecisive policy maker who may decide to change the structure of Ω (e.g. by shrinking this region) as the algorithm proceeds. Another weakness of the algorithm arises when the assumption $R \cap \Omega \neq \emptyset$ breaks down. Although Khatchian's original algorithm can easily identify $R \cap \Omega = \emptyset$ by not returning a feasible point after $4L(n + 1)^2$ iteration, the value L is difficult to estimate precisely in the above algorithm. However, it may be possible to determine some upper limit to L. Furthermore, changes in the structure of Ω seem to be in the nature of policy design. This would, at worst, increase p in (7.2), and hence L. It may also be possible to determine an upper limit to p at the beginning of the algorithm. Thus, the finiteness of the policy design process with the above algorithm can be demonstrated. In practice, experience dictates the choice of $\tilde{x}_0$ in Step 0. This initial value is chosen to be in a close neighbourhood of the region Ω. Thus, the speed at which $\underline{x}_c \in \Omega$ is attained with the above algorithm may be faster than the $4L\ (n+1)^2$ limit for Khatchian's algorithm.

8. Concluding Remarks

It is clear from the discussions in Sections 3 and 4 that given only preferred directions for key target variables and an element of consistency on the part of the policy maker (c.f. for example, Sen, 1970, p.63), it is possible to translate this 'qualitative' information into quantitative modification of the weights. In this sense, Frisch's preoccupation with efficient formulation of interview techniques in terms of different types of questions for the formulation of alternative preference functions seems slightly misplaced.

The most interesting extensions and applications of the methods presented above would seem to be in the important area of shadow price determination. Since there is a clear dual relationship between the weights of a preference function and the constraints, and hence the conventional multipliers, it is evident that by reformulating the methods we have discussed above, a direct application to the desirable determination of shadow prices would be possible. In the conventional literature on the determination of shadow prices, particularly with respect to the problem of economic development, there seems to be an imputation of an unwarranted halo of objectivity to such prices. The fact that they are, largely, as 'objective' as the weights of the preference function, is not always emphasized. Thus, as a 'truer' appreciation of the interlocking nature of weights, targets and instruments becomes evident due to the iterative nature of the method, an awareness of the objective constraints of the system also develops. Together with the weights, these latter are the principal determinants of shadow prices. It is, therefore, possible to apply a suitably formulated dual version of the above problem, and method, for the determination of shadow or accounting prices, and thus to demystify the somewhat excessive objectivity attached to these imputed values in the standard literature. Though this has been explicitly recognized in the excellent UNIDO (1972) exposition, no formal solution was provided (c.f. in particular, chapter 18 therein). It is our conjecture that the method presented above provides a formal solution to the problem of 'acceptable' shadow price determination.

On the other hand, there may be a temptation to interpret the above method as a solution to the problem of the historical revelation of preferences (inverse optimal control). This is very clearly an incorrect interpretation, as can be shown as follows.

The idea of "uncovering the objective function" once the decision has been made by others, is an old and illusive one. It has arisen independently in control theory and in economics. The interest from the latter area is due to the desire to reveal the past preferences of decision makers. The iterative method of Sections 3 and 4 may, in this case, be formulated so that the desired value $\underset{\sim}{x}^{d}$ is set according to the desirable historical conditions, the feasible region R is set to the model of the economy and the preferred value is set to be the actual historical value. The preferred value is fixed at the historical value throughout the procedure and is not changed, as would normally be done in Step 2 of the original procedure in Section 4. In this setting, the historical value which, in this case, is also the preferred value, has to be a feasible point since, by definition, the model of the economy must explain the historical event. Thus, for a linear model such as (3.4), the preferred value and all current optimal solutions are feasible and hence the $\underline{\delta}$ vector satisfies

$$N^T \underline{\delta} = \underline{0}.$$

Using this and (3.9), we can express the stepsize α in (5.1) (or α_z in (5.14)) as

$$\alpha = \frac{\mu < \underline{\delta}, Q_c(\underline{x}_c - \underline{x}^d) >}{< \underline{\delta}, Q_c \underline{\delta} > + \mu < Q_c \underline{\delta}, PH_c (Q_c \underline{\delta}_c) >}$$

$$= \frac{\mu < \underline{\delta}, N \underline{\lambda}_c >}{< \underline{\delta}, Q_c \underline{\delta} > + \mu < Q_c \underline{\delta}, PH_c (Q_c \underline{\delta}_c) >}$$

$$= 0.$$

Hence, if $\underline{\delta} \in R_o$, then the stepsize $\alpha = \alpha_z = 0$. This is not a serious limitation of the method in general since, if $\underline{x}_p$ is feasible ($\underline{x}_p \in R$), the decision maker can have exactly what he wants (i.e., since $\underline{x}_p \in \Omega \cap R$, setting $\underline{x}_n = \underline{x}_p$, the method stops). However, if the problem is to reveal the weighting matrix of a past decision, $\alpha = 0$ shows that the method is not suitable for this purpose.

Finally, consider the obvious choice for $\underline{x}_p$ which is $\underline{x}_p = \underline{x}^d$. However. this choice does not add any further information about the policy maker's local preferences. It is already known from the outset that the decision maker prefers $\underline{x}^d$ to any $\underline{x} \in R$. Hence $\underline{x}^d$ is preferred to $\underline{x}_c$ ($\in R$). The fact that $\underline{x}_p$, $\underline{x}_p \neq \underline{x}^d$, is preferred to $\underline{x}_c$, constitutes additional information about the local preferences of the policy maker. The uselessness of setting $\underline{x}_p = \underline{x}^d$ is supported by the above method. To show this, consider (5.1) and (3.9)

$$\underline{x}_n - \underline{x}_c = -\alpha(I - H_c N(N^T H_c N)^{-1} N^T)(\underline{x}_c - \underline{x}^d)$$

$$= -\alpha(I - H_c N(N^T H_c N)^{-1} N^T) H_c Q_c(\underline{x}_c - \underline{x}^d)$$

$$= -\alpha(I - H_c N(N^T H_c N)^{-1} N^T) H_c N \underline{\lambda}_c$$

$$= \underline{0}$$

which illustrates that this choice does not add any new information about the policy maker's preferences.

The main aim of the paper is to provide a method for the formalization of political preferences by means of the possible and necessary cooperation between policy makers and econometricians in the specific sense in which it was conceived by Frisch. We have, in this paper, attempted to provide a solution to a specific case dealt with in great detail by Frisch — but by relaxing some of the restrictive conditions and by considering in depth the problem of convergence. The extent to which the determination of the weights of the special case has been considered in relation to the almost arbitrary use to which such a form has been subject to in empirical and theoretical econometrics, may be an ex-post justification of Frisch's preoccupation with it — and a rationale for our exercise.

Finally, the similarity between Khatchian's ellipsoid algorithm, computing a feasible point of a given linearly constrained region, and the method in Section 4 highlights not only the complexity aspects of the method but also suggests that the search for an appropriate weighting matrix by the method is similar to the search of Khatchian's algorithm for a feasible point in $R \cap \Omega$.

FOOTNOTES

1 Also with a 'Nobel Prize' connotation, in that it was a lecture delivered at the invitation of the Federation of Swedish Industries, who initiated a tradition to invite Nobel Prize winners (in economics) to give a lecture on an optional subject.

2 Frisch used, interchangeably, the words 'econometrician', 'model builder', 'programming technician', 'scientist', etc. We will, therefore, retain the first of these terms to denote any one of these connotations.

3 i.e., the feasible set implied by, say, the behavioural and technical relations, constrained by the necessary accounting identities, characterizing an economy.

4 Here, again, Frisch used, interchangeably, 'policy maker', 'politician', 'economic planner', etc. We will keep to 'policy maker'.

5 "How to fix the bounds and how to determine the coefficients of the preference function are important practical problems. They are, indeed, so far-reaching that they lead us into a general consideration of the line of demarcation between the work of the politician and that of the scientist. Expressed briefly, and therefore necessarily without complete precision, we can say that the politician must introduce the human evaluations, the social value judgements, while the task of the scientist is objectively to find out what that factual situation is and what the inherent tendencies for change are, and what consequences COULD BE EXPECTED if one decided to put into effect such and such measures. In this work the scientist will simply have to take as data the goals themselves and the social value judgements at the back of them". (Frisch (1956), p.45).

6 In all cases, when we refer to 'optimal solutions' we refer, naturally, to the results obtained from a well-formulated constrained optimization problem - or, as Frisch would say, a Macroeconomic Programming problem. However, preserving the spirit in which Frisch conceived the problem, we should qualify it by stating that our prime concern, as econometricians cooperating with policy makers, will be to *optimally generate a politically acceptable path*.

7 "Many economists working on the theory and methods of economic planning have been in tune with Ragnar Frisch in stressing the optimization approach. *It appears somewhat surprising that so few have proceeded to taking up the question of how to establish a preference function.* (Johansen (1974), p.51: italics added.)

8 "The interview approach to the preference function is only a FIRST STAGE *in an iterative process which in each step proceeds by an optimal solution of the model".* (Frisch (1981), p.7: italics added.)

9 The distinction between decision and target variables is analogous to the two sets of variables characterising the final form of an econometric model. However, the analogy is not complete in the sense that target variables are a subset of the two sets of variables characterising the final form.

10 It should be noted that any positive semi-definite quadratic function

$$\arg\min\{ \langle \underline{a}, \underline{x} \rangle + \tfrac{1}{2} \langle \underline{x}, Q_c \underline{x} \rangle \mid \underline{x} \in R \} = \arg\min\{ q_c(\underline{x}) \mid \underline{x} \in R \}$$

for $Q_c \underline{x}^d = -\underline{a}$. Thus, provided this restriction on $\underline{x}^d$ is satisfied, arguments concerning the constrained minimum of $q_c(\underline{x})$ can also be extended to the more general quadratic function on the left of the above equality.

11 As the weighting matrix is symmetric, the quadratic functions (3.2) are said to penalize a deviation of the optimal solution of (3.2) from the desired value in a desirable direction as much as a deviation in an undesirable direction. However, the definition of the desired values implies that each element of $\underline{x}^d$ is assigned the best possible value — implying Pareto optimality — for that element. Hence, in this case, a deviation in a desirable direction of the actual solution from $\underline{x}^d$ can only imply a mis-specification of $\underline{x}^d$

12 This orthogonal decomposition is given by $N^T = [L \,\vdots\, 0]\,[Q]$, where L is an $m \times m$ lower triangular matrix (for $\dim(\underline{b}) = m$) and Q an orthogonal matrix $(Q\,Q^T = Q^T Q = I)$. Furthermore, Q is partitioned such that its first m rows are set to the $m \times n$ submatrix Q_1 and the last $n-m$ rows are denoted by Q_2 and Z is chosen to be Q_2^T.

13 In the presence of equality constraints, these constraints are used to eliminate some of the variables from the problem hence resulting in a problem with reduced dimensions (see Goldfarb and Todd, 1982). This, in turn, naturally reduces L in (7.2) and hence the bound on the number of iterations. This aspect is not considered further since it lies beyond the scope of this Section.

REFERENCES

Aspvall, B. and Stone, R.E. (1979) "Khatchian's Linear Programming Algorithm", STAN-CS-79-776, Stanford University.

Bland, R.G., Goldfarb, D. and Todd, M.J. (1981): The Ellipsoid Method: A Survey, Operations Research, V.29, N.6, pp.1039-1091.

Frisch, R. (1956) "Macroeconomics and Linear Programming", in Twentyfive Economic Essays in Honour of Eric Lindahl (pp. 38-67), Ekonomisk Tidskrift, Stockholm.

Frisch, R. (1959) "The Principle of Recurrant Planning" in Ekonomi, Politik, Samhalle. Festschrift for Bertil Ohlins Sixtieth Birthday, (pp. 67-74), Stockholm.

Frisch, R. (1961) "Numerical Determination of a Quadratic Preference Function for use in Macroeconomic Programming", in Studies in Honour of Gustavo del Vecchio, Padua.

Frisch, R. (1967) "A Multilateral Trade Crearing Agency", Economics of Planning, No. 2, pp. 97-105.

Frisch, R. (1972) "Cooperation Between Politicians and Econometricians on the Formalization of Political Preferences", University of Oslo, Institute of Economics Reprint Series, No.90.

Frisch, R. (1981) "From Utopian Theory to Practical Applications", American Economic Review, V.71, N.6, pp.1-16.

Gill, P.E. and Murray, W. (1974) "Newton-Type Methods for Linearly Constrained Optimization", in: Numerical Methods for Constrained Optimization, P.E. Gill and W. Murray (eds.), Academic Press, London.

Gill, P.E. and Murray, W. (1978) "Numerically Stable Methods for Quadratic Programming", Math. Programming, V.14, pp.349-372.

Goldfarb, D. (1969 "Extensions of Davidon's Variable Metric Method to Maximization Under Linear Inequality and Equality Constraints", SIAM J. Appl. Math., Vol.17, 739-764.

Goldfarb, D. and Todd, M.J. (1982) "Modifications and Implementation of the Ellipsoid Algorithm for Linear Programming", Math. Programming, V.23, N.1, pp.1-19.

Goodwin, R.M. and Velupillai, K. (1982) "Economic Systems and their Regulation", European University Institute Working Paper: 35, December, Florence.

Harsanyi, J.C. (1977) Rational Behaviour in Games and Social Situations, Cambridge University Press, Cambridge.

Householder, A.S. (1964) The Theory of Matrices in Numerical Analysis, Blaisdell, New York.

Johanssen, L. (1969) "An Examination of the Relevance of Kenneth Arrow's General Possibility Theorem for Economic Planning", Economics of Planning.

Johanssen, L. (1974) "Establishing Preference Functions for Macroeconomic Decision Models: Some Observations on Ragnar Frisch's Contributions", European Economic Review, V.5, pp. 41-66.

Johanssen, L. (1979) "Report on the Committee on Policy Optimization - UK" J. Econ. Dynamics and Control, V. 1, pp. 101-109.

Khatchian, L.G. (1979) "A Polynomial Algorithm in Linear Programming", Soviet Mathematics Doklady, V. 20, pp. 191-194.

Khatchian, L.G. (1980) "Polynomial Algorithms in Linear Programming", USSR Computational Mathematics and Mathematical Physics, V.20, No.1, pp.53-72.

Kozlov, M.K., Tarasov, S.P. and Khatchian, L.G. (1979) "Polynomial Solvability of Convex Quadratic Programming", Soviet Mathematics Doklady, V.20, pp.1108-1111.

Kornai, J. (1975) Mathematical Planning of Structural Decisions, North Holland, Amsterdam.

Luenberger, D. (1969) Optimization by Vector Space Methods, John Wiley, New York.

Ortega, J.M. and Rheinboldt, W.C. (1970) Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, London and New York.

Rustem, B., Velupillai, K. and Westcott, J.H. (1978) "Respecifying the Weighting Matrix of a Quadratic Objective Function", Automatica, Vol.14, pp.567-582.

Rustem, B. (1981) Projection Methods in Constrained Optimization and Applications to Optimal Policy Decisions, Springer-Verlag, Berlin.

Rustem, B. and Velupillai, K. (1979) "A New Approach to the Bargaining Problem", in: New Trends in Dynamic System Theory and Economics, M. Aoki and A. Marzollo (eds.), Academic Press, London and New York.

Sen, A. (1970) Collective Choice and Social Welfare, Holden-Day, San Francisco.

Sen, A. (1973) "Behaviour and the Concept of Preference", University of London. Inaugural Lecture at the London School of Economics.

Tinbergen, J. (1983) Review of: Theory of Economic Policy by A. Pagan and A. Preston, Kyklos, V. 36, N. 1, pp. 161-162.

UNIDO (1972) Guidelines for Project Evaluation, United Nations Publications, UNIDO, Vienna.

Von Wright, G.H. (1963) The Logic of Preference. Edinburgh University Press.

White, D.J. (1982) "A Selection of Multi-Objective Interactive Programming Methods", paper presented at the Conference on Multi-Objective Decision Making, University of Manchester.

EUI ECONOMICS DEPARTMENT WORKING PAPERS

No. 1 :	Jacques PELKMANS	The European Community and the Newly Industrialized Countries.
No. 3 :	Aldo RUSTICHINI	Seasonality in Eurodollar Interest Rates.
No. 9 :	Manfred E. STREIT	Essay on the Adequacy of an Abstraction.
No. 10 :	Kumaraswamy VELUPILLAI	When Workers Save and Invest: Some Kaldorian Dynamics.
No. 11 :	Kumaraswamy VELUPILLAI	A Neo-Cambridge Model of Income Distribution and Unemployment.
No. 12 :	Kumaraswamy VELUPILLAI/ Guglielmo CHIODI	On Lindahl's Theory of Distribution.
No. 22 :	Don PATINKIN	Paul A. Samuelson and Monetary Theory.
No. 23 :	Marcello DE CECCO	Inflation and Structural Change in the Euro-Dollar Market.
No. 24 :	Marcello DE CECCO	The Vicious/Virtuous Circle Debate in the '20s and the '70s.
No. 25 :	Manfred E. STREIT	Modelling, Managing and Monitoring Futures Trading: Frontiers of Analytical Inquiry.
No. 26 :	Domenico Mario NUTI	Economic Crisis in Eastern Europe - Prospects and Repercussions.
No. 34 :	Jean-Paul FITOUSSI	Modern Macroeconomic Theory An Overview.
No. 35 :	Richard M. GOODWIN/ Kumaraswamy VELUPILLAI	Economic Systems and their Regulation.
No. 46 :	Alessandra VENTURINI	Is the Bargaining Theory Still an Effective Framework of Analysis for Strike Patterns in Europe?.
No. 47 :	Richard A. GOODWIN	Schumpeter: the Man I Knew.
No. 48 :	J-P. FITOUSSI Daniel SZPIRO	Politique de l'Emploi et Réduction de la Durée du Travail.
No. 56 :	Berc RUSTEM Kumaraswamy VELUPILLAI	Preferences in Policy Optimization and Optimal Economic Policy.
No. 64 :	Marcello DE CECCO	Italian Monetary Policy in the 1980s.
No. 65 :	Giampaolo ROSSINI	Intraindustry Trade in Two Areas: Some Aspects of Trade Within and Outside a Custom Union.
No.66:	Wolfgang GEBAUER	Euromarkets and Monetary Control : the Deutschmark case.
No.67:	Gerd WEINRICH	On the Theory of Effective Demand under Stochastic Rationing.

No. 68:	Saul ESTRIN/ Derek C. JONES	The Effects of Worker Participation upon Productivity in French Producer Cooperatives.
No. 69:	Berc RUSTEM/ Kumaraswamy VELUPILLAI:	On the Formalization of Political Preferences : a Contribution to the Frischian Scheme.