MCPLIB: A Collection of Nonlinear Mixed Complementarity Problems

Steven P. Dirkse * Michael C. Ferris *

July 5, 1994

Abstract

The origins and some motivational details of a collection of nonlinear mixed complementarity problems are given. This collection serves two purposes. Firstly, it gives a uniform basis for testing currently available and new algorithms for mixed complementarity problems. Function and Jacobian evaluations for the resulting problems are provided via a GAMS interface, making thorough testing of algorithms on practical complementarity problems possible. Secondly, it gives examples of how to formulate many popular problem formats as mixed complementarity problems and how to describe the resulting problems in GAMS format. We demonstrate the ease and power of formulating practical models in the MCP format. Given these examples, it is hoped that this collection will grow to include many problems that test complementarity algorithms more fully.

The collection is available by anonymous ftp. Computational results using the PATH solver covering all of these problems are described.

1 Introduction

Recently, an extension to the GAMS modeling language has been developed which allows the formulation and solution of complementarity problems via GAMS. The use of GAMS speeds both the formulation of new models and the application of new algorithms to existing problems. As an aid to those developing new algorithms and to those wishing to formulate their own complementarity problems, we have developed a library of test problems. This report describes the origin and structure of the problems in the library. It is our intention

*Computer Sciences Department, University of Wisconsin, Madison, Wisconsin 53706. This material is based on research supported by the National Science Foundation Grant CCR–9157632 and the Air Force Office of Scientific Research Grant F49620–94–1–0036

that those developing complementarity solvers use the models in MCPLIB both to test their solvers and as a standard of comparison with other algorithms.

Several of the problems in the library have arisen from problems in economics. While an understanding of the underlying economics is not necessary in order to use the problems, it can be helpful; some of the problems have characteristics best understood in the context of the economics which determine them. Since this is the case, the economic background behind some of the models is given in this report; this has been done at a level which assumes little, if any, knowledge of economics.

While a GAMS model should be as self-documenting as possible, this report provides documentation which one could not hope to include with the code. It is hoped that by using this report, a user can gain a deeper understanding of the models in MCPLIB; references are provided as well.

Regardless of the origin of a complementarity problem, it must be correctly expressed as a mixed complementarity problem, or MCP, in order to be solved using GAMS. Letting $\overline{\mathbb{R}} := \{\mathbb{R}, -\infty, \infty\}$ denote the extended reals, we have the following:

Definition 1 (MCP) *Given a function $F : \mathbb{R}^n \to \mathbb{R}^n$ and bounds $l, u \in \overline{\mathbb{R}}^n$,*

$$\text{find} \quad x \in \mathbb{R}^n, \;\; u, v \in \mathbb{R}^n_+$$

$$\text{s.t.} \quad \begin{aligned} F(x) &= w - v && \text{(1a)} \\ l \le x &\le u && \text{(1b)} \\ (x - l)^\top w &= 0 && \text{(1c)} \\ (u - x)^\top v &= 0 && \text{(1d)} \end{aligned}$$

In contrast to the standard complementarity problem, lower and upper bounds on the variables x are explicitly included in MCP. This is of critical importance in developing efficient solution algorithms.

In Section 2, we describe some basic types of problems which serve as source problems for the models in the library. The relationship between the source problems and the MCP is summarized briefly and will be used in discussing the derivation of the models in the MCPLIB library. Included in the library are all the problems attempted in [13], [26], and [8]. Furthermore, new problem classes such as extended linear-quadratic programming and general equilibrium models are also included. In addition, a number of large general equilibrium models have been formulated by Rutherford [34] and are available directly from GAMS. The wide range of disciplines from which the MCPLIB models are drawn shows the versatility of the MCP format and the ease with which these models can be coded in GAMS. Currently, two solvers are available for solving these models, and new ones can easily be included. An AMPL version of the library, complete with solver interface routines, is currently under development.

Section 3 contains the descriptions of the larger, more complex models in the library, and a discussion of their derivation, where appropriate. The details of how to express these

MCP's in the GAMS language are not discussed in this paper, but the actual GAMS files are publicly available via anonymous ftp from `ftp.cs.wisc.edu:~/pub/mcplib/`. Section 4 contains numerical results for some of the problems in the library; these augment the numerical results given in [8].

A word about notation is in order. The transposition of a matrix or vector A is denoted by $A^\top$. The inner product of two vectors in $\mathbb{R}^n$ is defined as

$$\langle x, y\rangle := x^\top y = \sum_{i=1}^{n} x_i y_i$$

If β is a subset of $\{1, \ldots, n\}$, $x_\beta := (x_i), i \in \beta$. The concept of complementarity is central to our discussion. We will use the following notation to indicate a complementary function / variable pair and its associated bounds:

$$f(x) \geq 0, \qquad x \geq 0, \qquad \perp \tag{2}$$

This should be understood to mean that as well as satisfying the indicated constraints, $\langle f(x), x\rangle = 0$.

2 Problem Types

A number of well-known problem classes can be formulated as MCP's. The models in MCPLIB are drawn from nonlinear equations, nonlinear programming, nonlinear complementarity problems, and variational inequalities.

2.1 Nonlinear Equations

The nonlinear equations problem is that of finding a zero of a function $F : \mathbb{R}^n \to \mathbb{R}^n$, where the argument to F is unconstrained. If the bounds l and u in the MCP are set to $-\infty$ and ∞ respectively, the MCP variable x is unconstrained. Conditions (1c) and (1d) imply that both w and v are 0, so that (1a) reduces to requiring that x be a zero of F.

Nonlinear equations are of crucial importance in applications, and examples abound in the literature (e.g. the CUTE problems [3] and the Minpack-2 problems [1]). We include in MCPLIB examples of a distillation column model contributed by R. Fletcher and described in [23]. In this model, a steady state solution is sought in which a feed stream supplies material near the middle of a column and liquid and vapor are drawn out of the bottom and top of the column, respectively. GAMS models corresponding to each of three data sets (hydrocarbon-6, hydrocarbon-20, and methanol-8) are given. The damped Newton method employed by the PATH solver solves each of these problems. These problems are included as examples of how the many nonlinear equations models in the literature can be put into GAMS/MCP format.

2.2 Nonlinear Programming

Nonlinear programs consist of minimizing a smooth function of several variables over a feasible set defined by a number of constraints on these variables, as follows:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & x \in X := \{x \mid g(x) \leq 0, \quad x \geq 0\} \end{array} \tag{P}$$

Here $x \in \mathbb{R}^n$, while $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable functions. The *Karush–Kuhn–Tucker conditions* [20] for (P) are

$$\begin{array}{rll} \nabla f(x) + u^\top \nabla g(x) \geq 0, & x \geq 0, & \perp \\ -g(x) \geq 0, & u \geq 0, & \perp \end{array} \tag{KKT}$$

When f and g are convex functions, it is well known that solving (KKT) is sufficient for (P), in the sense that a solution (x, u) for (KKT) yields a solution x for (P). However, under slightly more restrictive assumptions, this equivalence can be made complete.

Theorem 2 ([20]) *Let f and g be convex, continuously differentiable functions defined on an open, nonempty subset X^0 of $\mathbb{R}^n$, and let g satisfy a suitable constraint qualification ([20]). Then $\bar{x}$ solves (P) if and only if there exists $\bar{u} \geq 0$ such that $(\bar{x}, \bar{u})$ solves (KKT).*

The simplest example of a nonlinear program is the quadratic program:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \frac{1}{2} x^\top Q x + c^\top x \\ \text{subject to} & Ax \leq b \end{array} \tag{QP}$$

Here $Q \in \mathbb{R}^{n \times n}$, $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$, with Q symmetric. If Q is positive semi-definite, the KKT conditions for (QP) are necessary and sufficient for a solution of (QP). Since (QP) does not bound x explicitly, its KKT conditions differ from those given for the problem (P):

$$\begin{array}{rll} Qx + c + A^\top u = 0, & x \text{ free}, & \perp \\ b - Ax \geq 0, & u \geq 0, & \perp \end{array}$$

These conditions constitute an MCP. If, in addition, $x \geq 0$, the problem has the form (KKT).

2.3 Nonlinear Complementarity Problems

Given a function $F : \mathbb{R}^n \to \mathbb{R}^n$ of x, the nonlinear complementarity problem (NCP) is to find x such that

$$F(x) \geq 0, \quad x \geq 0, \quad \perp . \tag{NCP}$$

The NCP is formulated as an MCP by setting $u = +\infty$ and $l = 0$. In this case, (1d) implies that $v = 0$, while the rest of (1) implies that $F(x)$ and x are non-negative and complementary. When F is affine, we have a linear complementarity problem (LCP).

A small example of an NCP, due to Kojima and Shindo [18], is defined by the polynomial function

$$F(x) := \begin{bmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_2^2 + x_1 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{bmatrix}. \tag{3}$$

This problem has two solution points,

$$x^1 = (\frac{\sqrt{6}}{2},\ 0,\ 0,\ 0.5), \quad x^2 = (1,\ 0,\ 3,\ 0),$$

and is difficult for simple Newton-type methods, since the LCP formed by linearizing F around $x = 0$ has no solution. Josephy [16] reports computational experience with a similar problem due to Kojima [17].

2.4 Variational Inequalities

An important and interesting problem, intimately related to the MCP, is the variational inequality, or VI: find $\bar{x} \in X$ such that

$$F(\bar{x})^\top (x - \bar{x}) \geq 0, \quad \forall\, x \in X \tag{VI}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ and $X \subset \mathbb{R}^n$ is convex. If the feasible set X in VI(F, X) is rectangular (i.e. $X := \{x \mid l \leq x \leq u\}$), then MCP and VI are completely equivalent, as their solution sets are identical. The proof of this is elementary. When X is polyhedral rather than rectangular, VI(F, X) can be reduced to an MCP by explicitly including the dual variables to the constraints defining X. Let $B := \{x \mid l \leq x \leq u\}$ and $X := \{x \mid Ax \leq b\}$, where $A \in \mathbb{R}^{m \times n}$. It can be shown that VI$(F, B \bigcap X)$ is equivalent to VI$(H, B \times \mathbb{R}^m_+)$, where

$$H(x, u) = \begin{bmatrix} F(x) + A^\top u \\ -Ax + b \end{bmatrix}.$$

When equality constraints are used to define X, the associated dual variables u are free.

3 The Model Library

The models discussed in this section have all been formulated in GAMS/MCP format. While many of the models are discussed in some detail, parameter values are not given in this report, since they can be found in the GAMS files. Table 1 lists the models currently contained in the library. In addition, some of the model types are described below.

Table 1: MCPLIB models

Model origin	GAMS file	Size
Nonlinear equations		
Distillation column modeling	hydroc20.gms	99
" "	hydroc06.gms	39
" "	methan08.gms	39
Nonlinear programming		
Quadratic programming	qp.gms	4
NLP test problem from Colville	colvncp.gms	15
Dual of Colville problem	colvdual.gms	20
Obstacle problems	obstacle.gms	N
Obstacle Bratu problems	bratu.gms	N
Nonlinear complementarity		
	josephy.gms	4
	kojshin.gms	4
Elastohydrodynamic lubrication	ehl_kost.gms	N
Variational inequalities		
Nash equilibrium	nash.gms	10
" "	choi.gms	14
Spatial price equilibrium	sppe.gms	27
" "	tobin.gms	42
Walrasian equilibrium	mathi*.gms	4
" "	scarfa*.gms	14
" "	scarfb*.gms	40
Traffic assignment	gafni.gms	5
Invariant capital stock	hanskoop.gms	14
Project Independence energy system (PIES)	pies.gms	42
Von Thünen land use	vonthun.gms	186
Extended linear-quadratic programming		
Optimal control	opt_cont.gms	N

3.1 Computing a Nash Equilibrium - `nash.gms`

The problem of computing a Nash equilibrium appears often in the literature (see [25, 12, 13]). The problem concerns a number of firms, each competitively producing a *common* good. We define the following:

N	number of firms, indexed $i = 1, \dots, N$
$x = (x_i)$	production vector; firm i produces a quantity x_i of the good
ξ	$e^\top x$, the sum total of the quantity being produced
$p(\xi)$	inverse demand function; $p(\xi)$ is the unit price at which consumers will demand (and actually purchase) a quantity ξ
$C_i(x_i)$	the production cost for firm i; note that this is the total cost, not a per–unit cost.

The firms comprise a market which we assume evolves over a number of time periods. At the beginning of each period, each firm sets its production level x_i so as to maximize its own profit, under the assumption that the production for all other firms remains constant at some level $x_j^*, j \neq i$. (These firms are said to operate in a *Nash manner.*) Intuitively, a Nash equilibrium point x^* is a production pattern in which *no* firm can increase its profit by unilaterally changing its level of production. Since no firm chooses to change its production in the current period, there is no change in the market, hence the equilibrium. Mathematically, a *Nash equilibrium* is a vector x^* such that

$$\forall\, i, \qquad x_i^* \in \arg\max_{x_i \geq 0} \qquad x_i\, p(x_i + \sum_{j \neq i} x_j^*) - C_i(x_i) \tag{4}$$

The KKT conditions for (4) take the following simple form:

$$\forall\, i, \qquad \nabla C_i(x_i) - p(\xi) - x_i\, \nabla p(\xi) \geq 0, \qquad x_i \geq 0, \qquad \perp \tag{NE}$$

which we call the Nash equilibrium conditions. In conformity with generally accepted economic behavior, the inverse demand function p is assumed to be strictly decreasing, the cost function C to be convex, and the "industry revenue curve" $\xi p(\xi)$ to be concave for $\xi \geq 0$. Under these assumptions, the objective function in (4) is concave [25]. By Theorem 2, the Nash equilibrium conditions (NE) are both necessary and sufficient for x^* to maximize (4). By combining the Nash equilibrium conditions for each i, we get an NCP in N variables.

The functions p and C used in the GAMS file `nash.gms` are defined below; c_i, L_i, β_i, and γ are parameters, with $\gamma > 1$.

$$p(\xi) = 5000^{\frac{1}{\gamma}}\, \xi^{\frac{-1}{\gamma}}$$

$$C_i(x_i) = c_i x_i \; + \; \frac{\beta_i}{1+\beta_i}\, L_i^{\frac{1}{\beta_i}}\, x_i^{\frac{\beta_i+1}{\beta_i}}$$

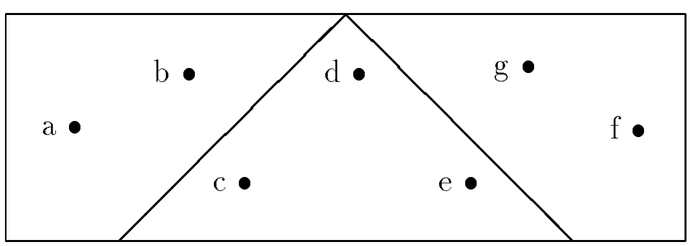
Another Nash equilibrium problem is given by Choi *et. al.* in [4]. In this problem, the firms are differentiated by the characteristics of the analgesic pain relievers they produce, rather than by production costs, while demand is determined by the prices and ingredient lists of the pain relievers. Data for this problem, and a description of the demand function, are given in the file `choi.gms`.

3.2 A Spatial Price Equilibrium Model - `sppe.gms`

In [11], Harker gives a number of models which describe the spatial and competitive structure of markets embedded in a network (i.e. a set of nodes and the arcs connecting them). Each node represents a unit or site separated spatially from the others. In each model, a *spatial price equilibrium* is sought. One competitive structure modeled is an oligopoly, a market situation in which a few producers control the deliveries to and demands from a large number of buyers. In our example, each producer tries to maximize the profit associated with his production of a single commodity common to all producers. We define the following:

L	set of distinct production units or *sites*
$W \subset L \times L$	set of transportation arcs between the sites in L
Q	set of producers, or *firms*, operating in the market
$I_q \in L$	set of sites controlled by firm $q \in Q$. The set of sites L is partitioned among the sets $I_q, q \in Q$.

Example 3 *Eight sites partitioned among 3 producers.*



$L = \{\mathrm{a}, \ldots, \mathrm{g}\}$
$Q = \{1, 2, 3\}$
$I_1 = \{\mathrm{a}, \mathrm{b}\}$
$I_2 = \{\mathrm{c}, \mathrm{d}, \mathrm{e}\}$
$I_3 = \{\mathrm{f}, \mathrm{g}\}$

$s_l, l \in L$	amount of commodity supplied (produced) by site l
$C_l(s_l)$	total cost of producing s_l units of output at site l (integral of inverse supply function)
$d_l, l \in L$	amount of commodity delivered (demanded) at site l
$\theta_l(d_l)$	purchase price dictated by the delivery to site l (inverse demand function)
$t_{ij}, ij \in W$	flow from site i to site j
$c_{ij}(t_{ij}), ij \in W$	unit transportation cost at level t_{ij}
d_{lq}	amount of commodity *produced by firm* q delivered to site l.

We will assume that each firm q acts in a Nash manner (see Section 3.1) when making decisions regarding the following quantities:

$s_i, i \in I_q$	the amounts produced at the sites q controls
$d_{lq}, l \in L$	amount of firm q's production delivered to each site in L
$t_{ij}, i \in I_q, j \in L$	flow from sites under firm q's control to each site in L.

The aggregation of these variables is firm q's strategy vector x_q. The constraints on x_q are those which ensure a conservation of flow at each site. Constraints for sites which firm q controls are more complicated than those for sites outside of firm q's control. The supply, delivery, and transportation variables are subject to lower and upper bounds, which we have taken to be 0 and $+\infty$, respectively. Thus, the set X_q of feasible strategies for the firm q is

$$X_q = \left\{ x_q := \begin{bmatrix} s_i \\ d_{lq} \\ t_{ij} \end{bmatrix} \geq 0 \;\middle|\; \begin{array}{lr} d_{lq} + \sum_{j \in L} t_{lj} = s_l + \sum_{i \in I_q} t_{il} \qquad (\forall\, l \in I_q) & \text{(5a)} \\ d_{lq} = \sum_{i \in I_q} t_{il} \qquad (\forall\, l \in L \setminus I_q) & \text{(5b)} \end{array} \right\}.$$

Let $X := \prod_{q \in Q} X_q$, so that $x \in X$ is a feasible strategy for all firms. Firm q's profit is then given by the function f_q:

$$f_q(x) \;:= \textstyle\sum_{l \in L} \theta_l(\sum_{j \in L} t_{jl}) d_{lq} - \sum_{i \in I_q} C_i(s_i) - \sum_{i \in I_q} \sum_{j \in L} c_{ij}(t_{ij}) t_{ij}, \tag{6}$$

so that firm q wishes to find a strategy x_q which solves the following problem:

$$\begin{array}{ll} \underset{x_q \in X_q}{\text{maximize}} & f_q(x) \\ \text{subject to} & x_p = \bar{x}_p \quad \forall\, p \neq q, \end{array} \tag{7}$$

where $\bar{x}_p$ is the current strategy employed by firm p. If we assume that, for all $l, i, j \in L$, $\theta_l(d_l)$ is a decreasing function, $C_l(s_l)$ is a convex function, and $c_{ij}(t_{ij})$ is an increasing function, then f_q is convex. If f_q is defined on the feasible set X and X contains a positive point, then, by applying a theorem from Rockafellar ([27], Theorem 27.4), we see that problem (7) is equivalent to $\mathrm{VI}(\nabla f_q, X_q)$, where f_q is differentiated with respect to x_q. A spatial price equilibrium [11] is therefore a point x which solves the following VI:

$$\begin{array}{ll} \text{find} & \bar{x} \in X \\ \text{s.t.} & \sum_{q \in Q} \nabla f_q(\bar{x})^\top (x_q - \bar{x}_q) \geq 0 \quad \forall\, x \in X \end{array} \tag{8}$$

The GAMS model for this problem can be obtained from (8) or, more directly, from the KKT conditions for (7). The particular model formulated contains 3 sites and 3 firms, so that each firm controls only one site; the relevant functions are defined as follows:

$$C_l(s_l) := \; \alpha_l s_l + \beta_l s_l^2, \qquad \theta_l(d_l) := \; \rho_l - \eta_l d_l, \qquad c_{ij}(t_{ij}) := \; \gamma_{ij} + \mu_{ij} t_{ij}^2.$$

While this particular example is somewhat limited, the GAMS model is coded for the general situation, where each firm controls multiple sites.

In [36], Tobin describes a spatial price equilibrium in a multi-commodity market modeled as a network. In this example, the variables are the prices at the various nodes in the network. These prices determine supply and demand, and not conversely, as in Harker's SPPE model. The competitive structure assumed in this example is one of perfect competition; it's "every node for itself". We define the following:

$l = 1, \dots, n$	the *nodes* (markets) in the network
$k = 1, \dots, p$	the commodities being traded in the network
$\pi = (\pi_{lk})$	price vector; for each node-commodity pair (l, k), π_{lk} is the unit price of commodity k at node l
$D_{lk}(\pi)$	demand for commodity k at node l
$S_{lk}(\pi)$	supply of commodity k at node l
$a = (ij)$	an arc in the network, from node i to node j
$A = [A_{la}]$	the standard node-arc incidence matrix. A is mainly zeros, with these exceptions: if $a = (ij)$, $A_{ia} = 1$ & $A_{ja} = -1$.
$t = (t_{ak})$	flow vector; for each arc-commodity pair (a, k), t_{ak} is the flow of commodity k on arc a
$c_{ak}(t_{ak})$	unit cost of tranportation service for commodity k on arc a

Section 2 of [36] gives the following conditions for a spatial price equilibrium (SPE):

Nonnegative flows, prices, demands, & supplies:

$$t_{ak} \geq 0, \quad \pi_{lk} \geq 0, \quad D_{lk} \geq 0, \quad S_{lk} \geq 0 \qquad \forall\, a, l, k \tag{9a}$$

Conservation of flow at each node:

$$S_{lk} + \sum_i t_{(il)k} = D_{lk} + \sum_j t_{(lj)k} \qquad \forall\, l, k \tag{9b}$$

Delivered price exceeds local price:

$$\pi_{ik} + c_{(ij)k}(t) \geq \pi_{jk} \qquad \forall\, a \; := (ij), k \tag{9c}$$

Delivered/local price difference or path flow = 0

$$\left\langle \pi_{ik} + c_{(ij)k} - \pi_{jk}, t_{ak} \right\rangle = 0 \qquad \forall\, a \; := (ij), k \tag{9d}$$

A set of flows and prices are feasible if they satisfy conditions (9a) and (9b). Condition (9c) and the complementarity condition (9d) imply that if the delivered price strictly exceeds the local price, no commodity is being delivered, and that if there is a commodity being delivered, its delivered price equals the local price.

If we relax the conservation of flow constraint (9b) to allow excessive supply, we get the following NCP:

$$c(t) + A^{\mathsf{T}}\pi \geq 0, \qquad t \geq 0, \qquad \perp \tag{10a}$$

$$S(\pi) - D(\pi) - At \geq 0, \qquad \pi \geq 0, \qquad \perp \tag{10b}$$

The following lemma gives conditions under which the conditions for a SPE are equivalent to the NCP defined in (10).

Lemma 4 ([9]) *Suppose the arc cost functions $c(t) > 0$ and the demand and supply functions are such that*

$$\pi_{lk} = 0 \;\Rightarrow\; D_{lk}(\pi) - S_{lk}(\pi) \geq 0 \tag{11}$$

Then a set of flows and prices $(\bar{t}, \bar{\pi})$ is a spatial price equilibrium iff it solves the NCP defined by (10a) - (10b).

In the GAMS model `tobin.gms`, the relevant functions are defined as follows:

$$c_{ak}(t) := \Gamma_{ak} + \Omega_{ak} t_{ak}^4 + \sum_{m \neq k} \Delta_{akm} t_{am}$$

$$S_{lk}(\pi) := \; B_{lk} + J_{lk}\pi_{lk}^2 + \sum_{i \neq l} u_{lik}\pi_{ik}$$

$$D_{lk}(\pi) := \; E_{lk} - G_{lk}\pi_{lk}^2 + \sum_{i \neq l} w_{lik}\pi_{ik}$$

3.3 A Walrasian Equilibrium Model - `mathi*.gms`

An equilibrium can be characterized as *Walrasian* if there are no goods for which demand strictly exceeds supply [37]. In [21], an economy containing a number of goods, a number of utility-maximizing consumers, and a number of profit-maximizing producers is described. Both consumers and producers act as price-takers, that is, they assume that the market price for each good does not change as a result of their actions. The role of the consumers here is to demand goods; this demand is determined by the prices. The producers determine their optimal levels of production based on these demands. Our objective is to find an equilibrium, or a steady state, for the economy. More specifically, we define the following:

$i = 1, \ldots, m$	indices corresponding to the m types of goods or commodities in the economy
$j = 1, \ldots, n$	index corresponding to the n *sectors* or types of production processes in the economy
$p = (p_i)$	vector of prices for the goods
$b = (b_i)$	vector of initial endowments for the goods (i.e. the amount of each good initially available)
$d(p) = (d_i(p))$	consumer demand functions; given a price vector, the demand for good i is $d_i(p)$
$y = (y_j)$	vector of activites; y_j is the activity or production level in sector j
$A = (a_{ij})$	technology matrix; a unit production level in sector j results in an output of a_{ij} units of good i. Negative values of a_{ij} indicate an input of good i is required for activity j. Column $A_{\cdot j}$ describes the process of sector j, while row of A_i indicates where good i is used and produced.

The equilibrium conditions ([35], Definition 5.1.3) are as follows:

No activity earns a positive profit:	$A^\top p \le 0$	(12a)
No good is in excess demand:	$b + Ay - d(p) \ge 0$	(12b)
No prices or activity levels are negative:	$p \ge 0 \quad y \ge 0$	(12c)
An activity earning a deficit is not run, and an operated activity runs at zero profit:	$y^\top A^\top p = 0$	(12d)
A good in excess supply has a zero price, and a positive price implies market clearance:	$p^\top(b + Ay - d(p)) = 0$	(12e)

At equilibrium, no activity earns a positive profit; if this were the case, others would step in to duplicate the activity, driving the profit to zero. Condition (12b) characterizes the equilibrium as Walrasian; there is no excess demand for any good. Condition (12e) implies that goods in excess supply have a zero price; if we assume that the goods are "desirable", (i.e. any good with a zero price must be in demand), then (12e) implies that all markets clear, or that supply equals demand.

A noteworthy property of Walrasian models is the assumption that the demand function $d(p)$ is homogeneous of degree 0 (i.e. $d(p) = d(tp) \quad \forall t > 0$). As a consequence, the equilibrium price vector is not unique; if p^* is an equilibrium price vector, so is tp^* for $t > 0$. An additional consequence of the homogeneity of d, shown in [21], is the singularity of the matrix $\nabla d(p)$. This singularity can make finding a solution difficult. Two customary ways of avoiding this singularity are normalizing the price vector or fixing one of the prices, called the numéraire price.

In the example given by Mathiesen [21], the consumer demand function $d(p)$ is determined by a single consumer; there is one production activity, and 3 goods. The problem is a difficult one because of the singularity of the Jacobian of the NCP formulation when no "fix" is applied, and because of the form of d:

$$d_i(\pi) := \frac{a_i \sum_k b_k \pi_k}{\pi_i}$$

If we require that $\sum_i a_i = 1$, then a_i determines the fraction of the budget $\sum_k b_k \pi_k$ spent on good i.

In [35], Scarf describes two similar Walrasian models, the smaller of which contains six commodities, eight activity sectors, and 6 consumers. Each consumer n has an initial asset e_{in} of each good i; the initial endowment b_i of good i is given by summing over all the consumers n. The individual initial assets are used in computing the demand function d, which is the sum of the individual consumers' demands. The equilibrium conditions (12) are the optimality conditions for this problem as well.

If α_{in} is the demand share parameter for good i and consumer n, and β_n is the elasticity of substitution for consumer n, then the demand function for this problem is

$$d_i(\pi) := \sum_n \alpha_{in} \pi_i^{\beta_n} \frac{\sum_k e_{kn} \pi_k}{\sum_k \alpha_{kn} \pi_k^{1-\beta_n}}$$

3.4 A Traffic Assignment Model - `gafni.gms`

In [2], a traffic assignment problem is given where there are 5 cities connected by a network of one–way links (see Figure 1). In each city i, there is a shipper who must ship d_i units of a commodity to city $(i+3)$. Thus, there are 5 origin-destination (OD) pairs in the network. There are only two paths or routes linking each OD pair, the inside and the outside paths. On each of these paths, a delay is incurred, which is equal to the sum of the delays on the links in that path. The delay on a link k is determined by the flow on and near link k, and is given in terms of a convex function g and a parameter $\gamma \geq 0$; we have taken $g(x) := 1 + x + x^2$. Figure 1 gives the configuration of the network, and the link delay functions. It is assumed that all flow not intended for a city will bypass that city.

Let x_i denote the amount shipped from city i via the outside path, and y_i the amount shipped via the inside path. Then the vectors $x = (x_i)$ and $y = (y_i)$ determine the flow on the paths, and also on each of the links. A flow is said to be feasible if

$$\begin{pmatrix} x \\ y \end{pmatrix} \in X := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \middle| \; x_i + y_i = d_i, \;\; x, y \geq 0 \right\}$$

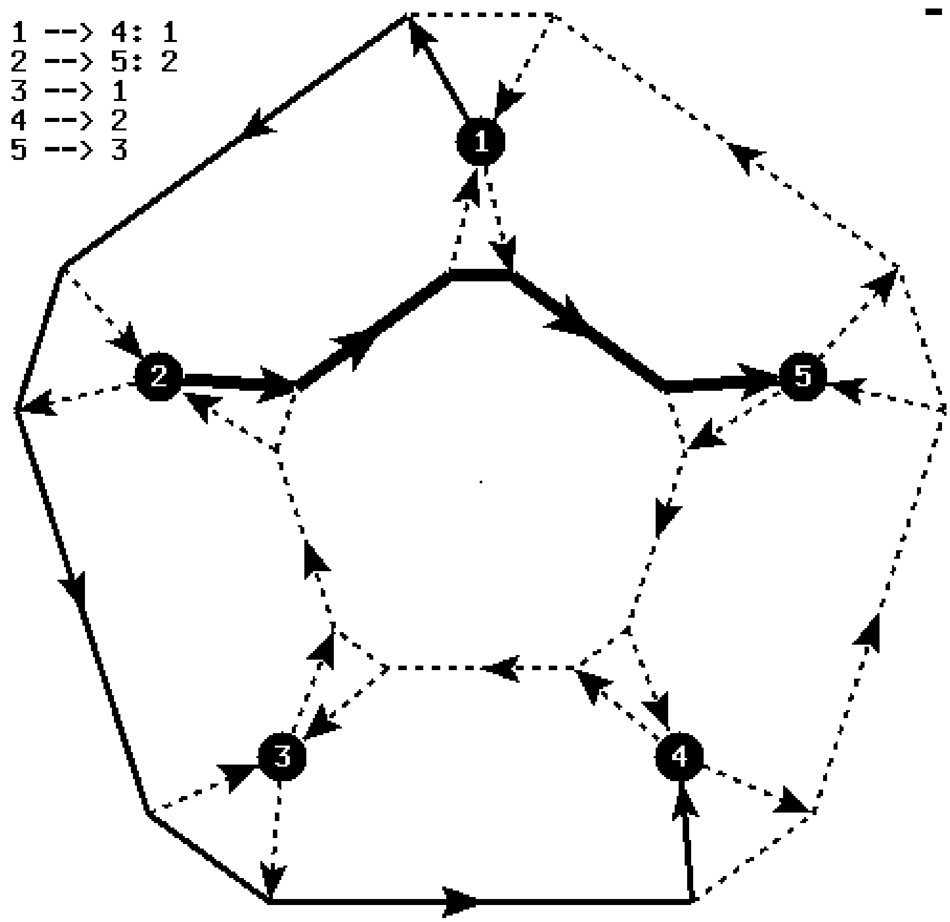
Given a flow $\begin{pmatrix} x \\ y \end{pmatrix}$, we define the *effective delay* between two cities in an OD pair to be the maximum delay among paths with *non-zero* flow between the two cities. The problem is to find a feasible flow in which each user has minimized her effective delay, subject to all other users' flows remaining constant. This occurs when the delay on every path with *non-zero* flow is the minimum among all paths between the corresponding OD pair. This flow is optimal in the sense that no user can reduce her effective delay by adjusting the flows she controls, while remaining feasible.

The conditions described in the above paragraph can be encapsulated by the optimality conditions VI(T, X), where

$$T \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} \text{outside-delay}(x) \\ \text{inside-delay}(y) \end{pmatrix}. \tag{13}$$

This VI in 10 variables and 5 demand constraints can be written simply as an NCP in 15 variables, if the demand constraints are relaxed to permit excess flow (there is no excess flow at the solution; clearly, sending excess flow increases any user's effective delay.) The simple demand constraints lead to NCP(G), where

$$G \begin{pmatrix} x \\ y \\ u \end{pmatrix} := \begin{pmatrix} \text{outside-delay}(x) - u \\ \text{inside-delay}(y) - u \\ x + y - d \end{pmatrix}.$$



highway links An arrow near midpoint indicates direction of flow. Delay on highway link k: $10g[flow_k] + 2\gamma g[flow_{\text{exit from } k}]$.

exit ramps An arrowhead indicates flow from a highway to a city. Delay on exit ramp k: $g[flow_k]$.

entrance ramps An arrowhead indicates flow from a city to a highway. Delay on exit ramp k: $g[flow_k] + \gamma g[flow_{\text{bypass of } k}]$.

bypass links No arrows; flow direction clear from figure. Delay on bypass link k: $g[flow_k]$.

Solid lines indicate positive flow.

Figure 1: Traffic Network

The problem can be expressed even more compactly by taking advantage of the constraint $x + y = d$ and the generality of the MCP model. Let $B := \{z \mid 0 \le z \le d\}$, then

$$X = \{a + Az \mid z \in B\}, \quad a = \begin{bmatrix} 0 \\ d \end{bmatrix}, \quad A = \begin{bmatrix} I \\ -I \end{bmatrix}.$$

Expressing VI(T, X) in term of z, we have the condition

$$\langle T(a + A\bar{z}), (a + Az) - (a + A\bar{z}) \rangle = \langle A^{\top} T(a + A\bar{z}), z - \bar{z} \rangle \ge 0 \quad \forall\, z \in B,$$

so that for $F(z) := A^{\top} T(a + Az)$, VI$(T, X)$ is equivalent to VI(F, B).

The intuition behind this latest VI is the clearest of any yet offered: $F_i(\bar{z})$ represents the difference in delay between the outside and inside paths from node i at optimality. When the difference is positive, the outside path is more expensive; all flow from node i should go to the inside. When the difference is negative, the inside path is more expensive; all flow from node i should go to the outside. When the difference is 0, any flow pattern from node i which satisfies the demand constraints is acceptable. Since the feasible set B is rectangular, the VI(F, B) is an MCP. Thus, we need only solve an MCP in 5 variables, rather than the forty-plus variables in the problem on the links, or the 15 variables in NCP(G).

3.5 Computing an Invariant Capital Stock - `hanskoop.gms`

Hansen and Koopmans [10] consider the problem of determining an invariant optimal capital stock. In this problem, an economy is assumed to grow over an infinite number of time periods. The technology (i.e. the production processes which can be run) and the available resources are assumed constant over all time periods. At the beginning of each time period, the economy invests its *capital goods* into the production processes, which produce both capital goods and *consumption goods*. The capital produced will be invested in the next period, while the consumption goods produced determine the utility of the investment. The total utility is a discounted sum; that is, the utility earned by an investment of capital at time t is discounted by a factor of α^t, where the discount factor $\alpha \in (0, 1)$. We wish to find an initial endowment of capital for which the investment strategy necessary to maximize the discounted sum of the utilities is constant. More formally, we have the following:

r	index for the set of resources types
i	index for the set of capital good types to be invested in production.
j	index for the set of production processes to run; each process consumes capital and resources, and produces capital and consumption goods.
$w = (w_r)$	The resources available at the beginning of each time period; this is assumed constant over time.
$z_t = (z_i)_t$	A *capital stock*; the amount of capital goods available for investment at the beginning of time period t.
$x_t = (x_j)_t$	The level at which to run the production processes during time period t. This effectively determines the investment of the capital stock z_t.

$v(x)$	Utility derived from the production/investment specified by x.
$A = (a_{ij})$	capital input matrix; running production process j at unit level requires a_{ij} units of capital good i ($A \geq 0$)
$B = (b_{ij})$	capital output matrix; running production process j at unit level produces b_{ij} units of capital good i ($B \geq 0$)
$C = (c_{rj})$	resource input matrix; running production process j at unit level requires c_{rj} units of resource good r ($C \geq 0$)
$0 < \alpha < 1$	discount factor for future utility

Assuming an integer time variable t, and given an initial capital stock z_0, we might wish to optimize our growth by solving the following:

$$\begin{array}{ll} \underset{x_t, z_t}{\text{maximize}} & \sum_{t=0}^{\infty} \alpha^t v(x_t) \\ \text{subject to} & \begin{array}{rcl} Ax_t & \leq & z_t \\ Bx_t & \geq & z_{t+1} \\ Cx_t & \leq & w \\ x_t & \geq & 0 \end{array} \end{array} \tag{14}$$

A solution of (14) maximizes the discounted sum of the utilities v; the feasibility conditions ensure that the *growth path* $\{(z_t, x_t)\}$ determining these utilities is consistent with the given technology and resource constraints. Notice that in (14), the initial capital stock z_0 is given; this stock determines the optimal growth path. Note also that the sequence of capital stocks $\{z_t\}$ is not fixed explicitly by the constraints in (14). However, it is likely that, over time, some optimal pattern of investment and return may evolve; that is, the growth path approaches a constant value.

This motivates the following problem. An initial capital stock z_0 is desired for which the optimal growth path does not vary. It should be noted that one cannot merely require that the path be constant, and optimize the choice of z_0. The invariance of the path must be a result of the optimality conditions in (14) and the choice of z_0, not of any explicit constraint. We will not derive here the conditions for a z_0 with a constant optimal growth path, since the motivation for the result is rather lengthy, and the proof is longer still. The interested reader is referred to [10], or to [7] for an example where v is linear.

We will assume that the utility function to be maximized in (14) is concave and continuously differentiable. Under some reasonable constraints on the technology, and a regularity condition on z_0, an initial capital stock z_0 whose optimal growth path (z_t, x_t) is constant satisfies the following NCP:

$$-\nabla v(x) + (A - \alpha B)^\intercal y + C^\intercal u \geq 0, \qquad x \geq 0, \qquad \perp \tag{15a}$$

$$(B - A)x \geq 0, \qquad y \geq 0, \qquad \perp \tag{15b}$$

$$-Cx + w \geq 0, \qquad u \geq 0, \qquad \perp \tag{15c}$$

A solution to NCP (15) suffices to determine an initial capital stock whose optimal growth path is constant; no regularity condition on z_0 is necessary in this direction. If $(\bar{x}, \bar{u}, \bar{y})$ satisfy (15), the capital stock $z_0 = A\bar{x}$.

3.6 Extended Linear-Quadratic Programming - `opt_cont.gms`

A number of recent papers have proposed an extended linear-quadratic programming (ELQP) model [29, 30] as a means of taking advantage of the special structure found in large-scale problems in multi-stage optimization [31], stochastic programming [32], and optimal control [29]. While problems formulated in this way are generally more difficult to solve than the conventional quadratic program, there exists an elegant duality theory for ELQP, which can be exploited in solution procedures. In this section, the ELQP is defined, and a significant special case is shown to be an instance of the MCP.

A problem in extended linear-quadratic programming is defined using the primal variables $u \in \mathbb{R}^n$, the dual variables $v \in \mathbb{R}^m$, and the nonempty, polyhedral sets $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$. Let $p \in \mathbb{R}^n$ and $P \in \mathbb{R}^{n \times n}$, and let $q \in \mathbb{R}^m$ and $Q \in \mathbb{R}^{m \times m}$, where Q and P are both symmetric positive semi-definite. In the ELQP model, some constraints are incorporated into a penalty or *monitoring function* added to the objective, rather than being considered explicitly. Given the set V and the matrix Q, this monitoring function is defined as

$$\rho_{VQ}(w) := \sup_{v \in V} \; w^\top v - \tfrac{1}{2} v^\top Q v \qquad \text{for } w \in \mathbb{R}^m \tag{16}$$

An extended linear-quadratic program may be defined using either a primal or dual form, both of which follow:

$$\underset{u \in U}{\text{minimize}} \qquad f(u) := p^\top u + \tfrac{1}{2} u^\top P u + \rho_{VQ}(q - Ru) \tag{P}$$

$$\underset{v \in V}{\text{maximize}} \qquad g(v) := q^\top v - \tfrac{1}{2} v^\top Q v - \rho_{UP}(R^\top v - p) \tag{D}$$

The difficulties in solving problems (P) and (D) arise from the monitoring functions ρ.

Theorem 5 ([28], Proposition 2.3) *The function ρ_{VQ} is lower semicontinuous, convex, and piecewise linear-quadratic: its effective domain*

$$\operatorname{dom} \rho_{VQ} := \{ w \in \mathbb{R}^m \mid \rho_{VQ}(w) < \infty \}$$

is a nonempty convex polyhedron that can be decomposed into finitely many polyhedral convex sets, on each of which ρ_{VQ} is quadratic (or linear); a similar result holds for ρ_{UP} and its effective domain.

Thus, the objective function f is convex and *piecewise* linear-quadratic, as is $-g$. This makes it difficult to apply techniques from smooth optimization in a straightforward manner.

However, duality theory can be used to show that problems (P) and (D) above are related through the following Lagrangian function:

$$L(u,v) := p^\top u + \tfrac{1}{2}u^\top Pu + q^\top v - \tfrac{1}{2}v^\top Qv - v^\top Ru, \tag{17}$$

with $f(u) = \sup_{v\in V} L(u,v)$ and $g(v) = \inf_{u\in U} L(u,v)$. The following theorem from Rockafellar [28] characterizes a pair of solutions to (P) and (D) as a saddle point of L.

Theorem 6 *It is always true that* $\inf(P) \geq \sup(D)$. *Furthermore, a pair* $(\bar{u},\bar{v})$ *is a saddle point of the Lagrangian* $L(u,v)$ *on* $U\times V$ *if and only* $\bar{u}$ *solves (P),* $\bar{v}$ *solves (D), and the optimum values are equal.*

The characterization of an optimal solution pair $(\bar{u},\bar{v})$ as a saddle point leads to a characterization in terms of a VI. We define

$$T\begin{pmatrix} u \\ v \end{pmatrix} := \begin{pmatrix} \nabla_u L(u,v) \\ -\nabla_v L(u,v) \end{pmatrix} = \begin{pmatrix} P & -R^\top \\ R & Q \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} p \\ -q \end{pmatrix} \tag{18}$$

and note from Theorem 6 that the pair $(\bar{u},\bar{v})$ is optimal for (P) and (D) if and only if $(\bar{u},\bar{v})$ solves VI$(T, U\times V)$.

Any ELQP can be reformulated as a conventional QP, and hence as a complementarity problem [32]. Unfortunately, this may greatly increase the problem size and disguise any special problem structure. Although specialized techniques can solve ELQP's quickly, we show that a frequently occurring special case of ELQP can be reformulated as an equivalent MCP, without any increase in size or loss of special structure. In a common practical situation [33, 32, 30], the feasible sets U and V are rectangular. In this case, the VI$(T, U\times V)$ defined by (18) is one involving only rectangular constraints, so that no reformulation is necessary to solve the problem as an MCP. In the remainder of this section, we discuss a continuous-time optimal control problem whose discretization results in a problem of this type.

Given a fixed time interval $[t_0, t_1]$, we define the primal problem in terms of the instantaneous control variables $u(t) \in U \subset \mathbb{R}^k$ and the left endpoint control variables $u_L \in U_L \subset \mathbb{R}^{k_L}$; the free state variables $x(t) \in \mathbb{R}^n$ depend on these control variables. The data for the problem (i.e. the matrices $\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}, \tilde{P}$, and $\tilde{Q}$, the vectors $\tilde{b}, \tilde{c}, \tilde{p}$, and $\tilde{q}$, and the feasible sets U and V) are generally assumed to vary continuously in t; we will assume that these matrices are constant as well. We seek to minimize the functional

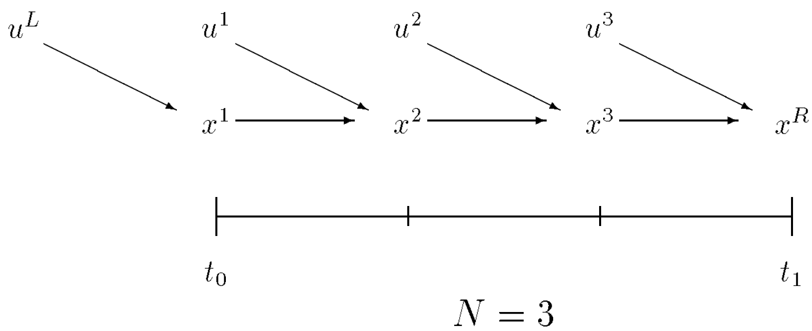
$$\begin{aligned}\mathcal{F}(u^L, u) := \int_{t_0}^{t_1} [\tilde{p}u(t) + \frac{1}{2}u(t)\tilde{P}u(t) - \tilde{c}x(t)]\,dt \quad + \quad p^L u^L + \frac{1}{2}u^L P_L u^L - c^R x(t_1) \\ + \int_{t_0}^{t_1} \rho_{V\tilde{Q}}(\tilde{q} - \tilde{C}x(t) - \tilde{D}u(t))\,dt \quad + \quad \rho_{V_R Q_R}(q^R - C_R x(t_1))\end{aligned}$$

over the state trajectory

$$\frac{dx}{dt}(t) = \tilde{A}x(t) + \tilde{B}u(t) + \tilde{b}, \qquad x(t_0) = B_L u^L + b^L, \tag{19}$$

where the subscripts L and R denote data and variables used to define boundary conditions at the left and right endpoints, respectively. In this model, the feasible sets U, U_L, V, and V_R are bounded rectangular sets.

The ELQP model arises as a discretization of the continuous problem above. The interval $[t_0, t_1]$ is divided into N segments, so that the variables $u(t)$ and $x(t)$ are discretized as follows,



where the arrows indicate the dependence of the state variables on previous states and controls, as determined by (19). If we assume that $t_1 - t_0 = 1$, the resulting discrete-time ELQP is that of minimizing

$$\frac{1}{N}\sum_{1}^{N}[\tilde{p}u^i + \frac{1}{2}u^i\tilde{P}u^i - \tilde{c}x^i] \quad + \quad p^L u^L + \frac{1}{2}u^L P^L u^L - c^R x^R$$
$$+\frac{1}{N}\sum_{1}^{N}\rho_{V\tilde{Q}}(\tilde{q} - \tilde{C}x^i - \tilde{D}u^i) \quad + \quad \rho_{V_R Q_R}(q^R - C^R x^R)$$

subject to the state constraints

$$x^1 \;=\; B_L u^L + b^L \tag{20}$$
$$x^{i+1} \;=\; x^i + \frac{1}{N}(\tilde{B}u^i + \tilde{A}x^i + \tilde{b}) \quad i = 1, \ldots, N-1 \tag{21}$$
$$x^R \;=\; x^N + \frac{1}{N}(\tilde{B}u^N + \tilde{A}x^N + \tilde{b}). \tag{22}$$

If we define $A := I + \frac{1}{N}\tilde{A}$, $B := \frac{1}{N}\tilde{B}$, $b := \frac{1}{N}\tilde{b}$, $C := \frac{1}{N}\tilde{C}$, $c := \frac{1}{N}\tilde{c}$, $D := \frac{1}{N}\tilde{D}$, $P := \frac{1}{N}\tilde{P}$,

$p := \frac{1}{N}\tilde{p}$, $Q := \frac{1}{N}\tilde{Q}$, and $q := \frac{1}{N}\tilde{q}$, we obtain the following ELQP:

$$\begin{aligned} \underset{u^L,u^i,x^i,x^R}{\text{minimize}} \quad & \mathcal{F}_D(u^L,u^i,x^i,x^R) := \\ & \sum_1^N [pu^i + \frac{1}{2}u^i P u^i - cx^i] \;+\; p^L u^L + \frac{1}{2}u^L P_L u^L - c^R x^R \\ & + \sum_1^N \rho_{VQ}(q - Cx^i - Du^i) \;+\; \rho_{V_R Q_R}(q^R - C_R x^R) \end{aligned}$$

subject to the constraints

$$\begin{aligned} x^1 &= B_L u^L + b^L \\ x^{i+1} &= Bu^i + Ax^i + b \quad i = 1,\ldots,N-1 \\ x^R &= Bu^N + Ax^N + b. \end{aligned}$$

Using (18), we can express the optimality conditions for the discrete-time minimization problem as the VI($F, U_L \times U^N \times \mathbb{R}^{n(N+1)} \times V^N \times V_R \times \mathbb{R}^{n(N+1)}$), with

$$F\begin{pmatrix} u \\ x \\ v \\ y \end{pmatrix} = \begin{bmatrix} \bar{P} & 0 & -\bar{D}^\top & -\bar{B}^\top \\ 0 & 0 & -\bar{C}^\top & I - \bar{A}^\top \\ \bar{D} & \bar{C} & \bar{Q} & 0 \\ \bar{B} & \bar{A} - I & 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ x \\ v \\ y \end{bmatrix} + \begin{bmatrix} \bar{p} \\ -\bar{c} \\ -\bar{q} \\ \bar{b} \end{bmatrix},$$

where

$$\bar{P} := \begin{bmatrix} P_L & & & \\ & P & & \\ & & \ddots & \\ & & & P \end{bmatrix}, \quad \bar{D} := \begin{bmatrix} 0 & D & & \\ & 0 & \ddots & \\ & & \ddots & D \\ & & & 0 \end{bmatrix}, \quad \bar{B} := \begin{bmatrix} B_L & & & \\ & B & & \\ & & \ddots & \\ & & & B \end{bmatrix},$$

$$\bar{C} := \begin{bmatrix} C & & & \\ & \ddots & & \\ & & C & \\ & & & C_R \end{bmatrix}, \quad \bar{A} := \begin{bmatrix} 0 & & & \\ A & 0 & & \\ & \ddots & \ddots & \\ & & A & 0 \end{bmatrix}, \quad \bar{Q} := \begin{bmatrix} Q & & & \\ & \ddots & & \\ & & Q & \\ & & & Q_R \end{bmatrix},$$

$$\bar{p} := \begin{bmatrix} p^L \\ p \\ \vdots \\ p \end{bmatrix}, \quad \bar{c} := \begin{bmatrix} c \\ \vdots \\ c \\ c^R \end{bmatrix}, \quad \bar{q} := \begin{bmatrix} q \\ \vdots \\ q \\ q^R \end{bmatrix}, \quad \bar{b} := \begin{bmatrix} b^L \\ b \\ \vdots \\ b \end{bmatrix},$$

and the dots represent replication N times.

In the GAMS implementation, the data elements for the continous-time problem are generated randomly, where the matrices $\tilde{P}$ and $\tilde{Q}$ are generated to be positive (semi)definite.

The division by N takes place during the formation of the discretized problem. Note that the discrete-time problem makes use of the function $\mathcal{F}_D(u^L, u^i, x^i, x^R)$ in the variables u and x, while the continuous problem is expressed as a minimization over u only. While it is possible to express the discrete time problem without using the x variables, this results in a dense problem. For this reason, the state variables x and y are retained in the MCP formulation.

3.7 An Obstacle Problem - `obstacle.gms`

The obstacle problem [5] consists of finding the equilibrium position of an elastic membrane subject to a vertical force f pushing upwards. In our example, we consider a membrane with height v on a domain $\mathcal{D} := (0,1) \times (0,1)$. We restrict our attention to those functions v in the space $H_0^1(\mathcal{D})$ of functions with compact support in $\mathcal{D}$ such that v and $\|\nabla v\|^2$ belong to the square integrable class $L^2(\mathcal{D})$. Note that this implies that $v = 0$ on the boundary of $\mathcal{D}$. In addition, we have lower and upper bounds v_ℓ and v_u on v which represent the position of solid objects below and above the membrane, respectively. The membrane's equilibrium position is its position of minimum energy, where the energy of the membrane is given by the quadratic functional $q(v)$ in the following quadratic program:

$$\begin{array}{ll} \underset{v}{\text{minimize}} & q(v) = \frac{1}{2}\int_{\mathcal{D}} \|\nabla v\|^2 \, d\mathcal{D} - \int_{\mathcal{D}} f v d\mathcal{D} \\ \text{subject to} & v \in H_0^1(\mathcal{D}) : v_\ell \leq v \leq v_u \end{array}. \tag{23}$$

In [24], the force f is taken to be the constant $c = 1$.

In order to solve this problem numerically, the domain $\mathcal{D}$ is discretized by a triangulation of a rectangular grid with grid spacing $h := \frac{1}{N+1}$ in both the X and Y axes. The function v is then approximated by a piecewise linear function which can be represented by its values $v_{i,j}$, for $i, j = 1, \ldots, N$, at the N^2 interior vertices of the triangulation. Using this approximation, the objective function q in (23) can be reduced (see for example [24]) to a quadratic function

$$q(v) := \tfrac{1}{2} v^\top M v - q^\top v, \tag{24}$$

where the components of $v \in \mathbb{R}^{N^2}$ are the values $v_{i,j}$ at the vertices of the triangularization, $q_{i,j} = ch^2$, and M is the usual pentadiagonal matrix obtained via a difference approximation of the Laplacian operator (diagonal entries of 4, off-diagonal entries of -1). Given the constraints $v_\ell \leq v \leq v_u$, the optimality conditions for minimizing the discretized $q(\cdot)$ can be written as the following MCP:

$$f(v) := Mv + q \ \text{ free}, \qquad v_\ell \leq v \leq v_u, \qquad \perp . \tag{25}$$

If the force f acting on the membrane is taken taken to be the nonlinear function λe^v, the *obstacle Bratu* problem results. This problem, solved in [22, 15], differs from the one just described in that the components of the vector q are no longer constant but are a function of v, i.e., $q_{i,j} = \lambda e^{v_{i,j}}$.

3.8 The Elastohydrodynamic Lubrication Problem - `ehl_kost.gms`

The problem of the elastohydrodynamic lubrication of cylinders in line contact is considered by Kostreva [19]. A particular example would consider (cylindrical) roller bearings lubricated by oil. The standard mathematical model for this problem is governed by 3 equations: a linear integral equation for the deformation of the cylinders, Reynolds' differential equation for the pressure in the lubricant, and a linear integral equation which represents a balance of load constraint. If the lubricant pressure at position x is represented by $p(x)$, then the thickness h of the lubricant film between the cylinders at position x is given by

$$h(x) = x^2 + k - \frac{2}{\pi}\int_a^b p(s)\ln|x-s|ds, \tag{26}$$

where k is a free variable of the model, x_a is an inlet point and x_b is an outlet point to be determined from the model solution, with $x_a < x_b$. The pressure will be positive between the inlet and outlet points, while the boundary conditions are $p(x_a) = p(x_b) = p'(x_b) = 0$. In the region of positive pressure, Reynolds' equation, which relates lubricant pressure to lubricant film thickness, holds:

$$R(p,k) := -\frac{d}{dx}\left(\frac{h(x)^3}{e^{\alpha p}}\frac{dp}{dx}\right) + \lambda\frac{dh}{dx} = 0. \tag{27}$$

Downstream of x_b, the pressure will be 0, so that Reynolds' equation need not be satisfied; in this area, $R(p,k)$ is allowed to become positive and reduces to $\lambda\frac{dh}{dx}$. Since $\lambda > 0$, this represents a divergence of the cylinders downstream of the outlet point. The final equation represents a constraint placed on the cumulative pressure required by the specified load on the cylinders:

$$T(p,k) := 1 - \frac{2}{\pi}\int_a^b p(s)ds = 0. \tag{28}$$

Given the inlet point x_a, the complementarity form of this problem makes use of finite difference approximations to R and T on the interval $[x_a, x_F]$, where x_F is chosen to be far downstream, so that $x_F > x_b$. Given a uniform grid of N intervals such that $x_F = x_a + N\Delta x$, let $p_i = p(x_a + i\Delta x)$ and let $h_j = h(x_a + j\Delta x)$ for $i = 1, \ldots, N,\ j = i \pm \frac{1}{2}$. The values of h_j at the intermediate points can be approximated by numerical integration of (26) or by the following, computationally recommended, integral obtained from (26) via integration by parts:

$$h(x) = x^2 + k + 1 + \frac{2}{\pi}\int_{x_a}^{x_b}(s-x)\ln|x-s|\left(\frac{dp}{ds}\right)ds. \tag{29}$$

In the GAMS model, both h_j and T are approximated using the trapezoidal rule. The formula for h_j is substituted into the finite difference approximation to Reynolds' equation at the points x_i for $i = 1, \ldots, N$ as follows:

$$R_i(k,p) := -\frac{1}{(\Delta x)^2}\left[\frac{(h_{i+\frac{1}{2}})^3}{\exp(\alpha p_{i+\frac{1}{2}})}(p_{i+1}-p_i) - \frac{(h_{i-\frac{1}{2}})^3}{\exp(\alpha p_{i-\frac{1}{2}})}(p_i - p_{i-1})\right] + \frac{\lambda}{\Delta x}(h_{i+\frac{1}{2}} - h_{i-\frac{1}{2}}). \quad (30)$$

The final MCP is given by

$$T(k,p) = 0, \qquad k \text{ free}, \qquad \perp \quad (31a)$$
$$R_i(k,p) \geq 0, \qquad p_i \geq 0, \qquad \perp, \quad \text{for } i = 1, \ldots, N. \quad (31b)$$

As mentioned earlier, the location of the free boundary x_b is not known *a priori*; it is determined as part of the solution to the complementarity problem. This is in contrast to other methods proposed for this problem, which rely on heuristics to locate the free boundary. In [19], Kostreva considers examples where the free boundary has been mislocated by these techniques, as well as other examples where the computed film thickness h differs from previous results.

The elastohydrodynamic lubrication model is interesting both because of its highly nonlinear nature and because of its potentially large size. Unfortunately, it is a fully dense, so that sparse techniques cannot be used to improve performance. In his computational work, Kostreva [19] used a grid of size 0.05 on an interval of length 5, resulting in a highly nonlinear model with 100 equations. However, for higher pressure and load conditions, the solution to this problem develops a large pressure spike, which can be difficult to compute, and leads to finer grid approximations and larger problems.

4 Numerical Results

In this section, we give numerical results obtained by solving some of the models described above. Unless otherwise indicated, these results were obtained through the use of the PATH solver for MCP, described in [8] and running as a GAMS subsystem on a DECstation 5000/125. Solution times given are those reported as the resource usage in the GAMS listing file. Computational results for the models not considered in this section, and comparisons of the PATH solver to other MCP algorithms, are found in [8].

We consider first the optimal control problem described in Section 3.6. This problem can be expressed and solved as both an MCP or a QP; we have taken both approaches in solving this problem. In our computational tests, we have solved a single continuous time problem with 8 control and 8 state variables and 8 dual control and 8 dual state variables. By varying the number of points N in the discretization of the continuous interval, we vary the problem size. The table below shows the times required to solve the problem for different values of N. The MCP's were solved using the PATH solver, while the QP's were solved using GAMS/MINOS 5.3 with the default parameters. The solution times and pivot counts were obtained by averaging the results of several runs using different random number seeds. A time limit of 10 hours was placed on all the runs, as the larger problems were not solvable using MINOS.

Table 2: Solution Times - Optimal Control Model

	MCP				QP	
N	size	nonzeros	pivots	time (sec)	size	time
15	512×512	8448	220	12	257×641	54
31	1024×1024	17152	432	45	513×1281	953
127	4096×4096	69376	1828	717	2049×5121	28423
255	8192×8192	139008	3967	3550	na	na
350	11232×11232	190687	5549	7417	na	na

Table 2 illustrates the effectiveness of the PATH solver in solving large complementarity problems, and also provides further evidence for the validity of the MCP model. In the case of the ELQP given in Section 3.6, the QP formulation has proven much more difficult to solve than an equivalent formulation as an MCP.

Table 3: Solution Times - Obstacle Model A

	MCP				
N	v_0	size	nonzeros	pivots	time (sec)
75	ℓ	5625×5625	28124	2123	544
75	e	5625×5625	28124	3505	2713

Table 4: Solution Times - Obstacle Model B

	MCP				
N	v_0	size	nonzeros	pivots	time (sec)
75	ℓ	5625×5625	28124	6367	2692
75	u	5625×5625	28124	4885	1623
75	$\frac{(\ell+u)}{2}$	5625×5625	28124	1455	1202

Table 5: Solution Times - Obstacle Model C

	MCP				
N	v_0	size	nonzeros	pivots	time (sec)
75	ℓ	5625×5625	28124	6205	3073
75	u	5625×5625	28124	5047	1850
75	$\frac{(\ell+u)}{2}$	5625×5625	28124	1942	1782

The MCP arising from the obstacle problem considered in Section 3.7 was solved using the PATH solver for $N = 75$ and with the obstacles A, B, and C, where the lower and upper bounds for obstacle A are

$$v_\ell(x, y) = \sin(3.2x)\sin(3.3x), \quad v_u(x, y) = 2000,$$

for obstacle B,

$$v_\ell(x, y) = (\sin(9.2x)\sin(9.3x))^3, \quad v_u(x, y) = (\sin(9.2x)\sin(9.3x))^2 + .02,$$

and for obstacle C,

$$v_\ell(x, y) = (16x(1-x)y(1-y))^3, \quad v_u(x, y) = (16x(1-x)y(1-y))^2 + .01.$$

The data in Tables 2, 3, 4, and 5 indicate that the PATH solver performs a large number of pivot steps when solving these large problems. This is to be expected: the pivotal techniques employed by the PATH solver place it among those QP solvers which use an active set strategy. For solvers that add or subtract one constraint at a time from the active set, the number of pivots required is bounded below by the difference in size between the optimal and initial set of active constraints. This bound can be expected to grow with the size of the problem, as is seen in the computational examples presented in Table 2.

Table 6: Solution Times - EHL Model

N	α	λ	p_0	major	pivots	time (sec)
100	2.832	6.057	hertz	6	89	20
100	3.746	9.889	hertz	21	927	98
100	4.477	9.692	hertz	13	381	52

The nonlinear nature of the elastohydrodynamic lubrication model makes it particularly amenable to solution by the PATH solver. The stabilization techniques used by this solver enable the solution of models representing high load and speed with a minimum amount of dependence on the starting point used. Table 6 gives the solution times, major iterations and total pivots used to solve the EHL model for the indicated values of the parameters α and λ. The parameter values represent increasing load and speed conditions for the bearing being modeled. The starting points were all taken to be the solution to the Hertzian (dry) case. It was not necessary to use solution points for lower values of α and λ as initial points when solving for higher parameter values, as was done by Kostreva in [19].

References

[1] B. M. Averick, R. G. Carter, and J. J. More. The MINPACK-2 test problem collection (preliminary version). Technical Report MCS-TM-150, Argonne National Laboratory, Argonne, USA, May 1991.

[2] D. P. Bertsekas and E. M. Gafni. Projection methods for variational inequalities with application to the traffic assignment problem. *Mathematical Programming Study*, 17:139–159, 1982.

[3] I. Bongartz, A. R. Conn, N. Gould, and P. Toint. CUTE: Constrained and unconstrained testing environment. Publications du Départment de Mathématique Report 93/10, Facultés Universitaires De Namur, 1993.

[4] S. C. Choi, W. S. DeSarbo, and P. T. Harker. Product positioning under price competition. *Management Science*, 36(2):175–199, 1990.

[5] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North-Holland, New York, 1978.

[6] A. R. Colville. A comparative study on nonlinear programming codes. Technical Report 320–2949, IBM New York Scientific Center, June 1968.

[7] R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem.* Academic Press, New York, 1992.

[8] S. P. Dirkse and M. C. Ferris. The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems. *Optimization Methods & Software*, 1994. To appear.

[9] T. L. Friesz, R. L. Tobin, T. E. Smith, and P. T. Harker. A nonlinear complementarity formulation and solution procedure for the general derived demand network equilibrium problem. *Journal of Regional Science*, 23:337–359, 1983.

[10] T. Hansen and T. C. Koopmans. On the definition and computation of a capital stock invariant under optimization. *Journal of Economic Theory*, 5:487–523, 1972.

[11] P. T. Harker. Alternative models of spatial competition. *Operations Research*, 34:410–425, 1986.

[12] P. T. Harker. Accelerating the convergence of the diagonalization and projection algorithms for finite-dimensional variational inequalities. *Mathematical Programming*, 41:29–59, 1988.

[13] P. T. Harker and B. Xiao. Newton's method for the nonlinear complementarity problem: A B–differentiable equation approach. *Mathematical Programming*, 48:339–357, 1990.

[14] W. W. Hogan. Energy policy models for Project Independence. *Computers & Operations Research*, 2:251–271, 1975.

[15] R. H. W. Hoppe and H. D. Mittelmann. A multi-grid continuation strategy for parameter-dependent variational inequalities. *Journal of Computational and Applied Mathematics*, 26:35–46, 1989.

[16] N. H. Josephy. Newton's method for generalized equations. Technical Summary Report 1965, Mathematics Research Center, University of Wisconsin–Madison, Madison, Wisconsin, 1979.

[17] M. Kojima. Computational methods for solving the nonlinear complementarity problem. *Keio Engineering Reports*, 27:1–41, 1974.

[18] M. Kojima and S. Shindo. Extensions of Newton and quasi-Newton methods to systems of PC^1 equations. *Journal of Operations Research Society of Japan*, 29:352–374, 1986.

[19] M. M. Kostreva. Elasto-hydrodynamic lubrication: A non-linear complementarity problem. *International Journal for Numerical Methods in Fluids*, 4:377–397, 1984.

[20] O. L. Mangasarian. *Nonlinear Programming*. McGraw–Hill, New York, 1969.

[21] L. Mathiesen. An algorithm based on a sequence of linear complementarity problems applied to a Walrasian equilibrium model: An example. *Mathematical Programming*, 37:1–18, 1987.

[22] E. Miersemann and H. D. Mittlemann. Continuation for parametrized nonlinear variational inequalities. *Journal of Computational and Applied Mathematics*, 26:23–34, 1989.

[23] J. J. Moré. A collection of nonlinear model problems. Technical Report MCS-P60-0289, Argonne National Laboratory, Argonne, USA, February 1989.

[24] J. J. Morè and G. Toraldo. On the solution of large quadratic programming problems with bound constraints. *SIAM Journal on Optimization*, 1:93–113, 1991.

[25] F. H. Murphy, H. D. Sherali, and A. L. Soyster. A mathematical programming approach for determining oligopolistic market equilibrium. *Mathematical Programming*, 24:92–106, 1982.

[26] J.-S. Pang and S. A. Gabriel. NE/SQP: A robust algorithm for the nonlinear complementarity problem. *Mathematical Programming*, 60(3):295–338, 1993.

[27] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.

[28] R. T. Rockafellar. Linear-quadratic programming and optimal control. *SIAM Journal on Control and Optimization*, 25(3):781–814, 1987.

[29] R. T. Rockafellar. Multistage convex programming and discrete-time optimal control. *Control and Cybernetics*, 17(2-3):225–245, 1988.

[30] R. T. Rockafellar. Computational schemes for large-scale problems in extended linear-quadratic programming. *Mathematical Programming*, 48:447–474, 1990.

[31] R. T. Rockafellar. Large-scale extended linear-quadratic programming and multistage optimization. In S. Gomez, J. P. Hennart, and R. A. Tapia, editors, *Advances in Numerical Partial Differential Equations and Optimization*, chapter 15, pages 247–261. SIAM Publications, 1991.

[32] R. T. Rockafellar and R. J.-B. Wets. A Lagrangian finite generation technique for solving linear-quadratic problems in stochastic programming. *Mathematical Programming Study*, 28:63–93, 1986.

[33] R. T. Rockafellar and R. J.-B. Wets. Linear-quadratic programming problems with stochastic penalties: the finite generation algorithm. In V. I. Arkin, A. Shiraer, and R. J.-B. Wets, editors, *Springer Verlag Lecture Notes in Control and Information Sciences*, IIASA Series 81, pages 545–560. Springer Verlag, 1986.

[34] T. F. Rutherford. Applied general equilibrium modeling with MPSGE as a GAMS subsystem. Manuscript, Department of Economics, University of Colorado, Boulder, Jan. 1994.

[35] H. E. Scarf. *The Computation of Economic Equilibria*. Yale University Press, New Haven, Conneticut, 1973.

[36] R. L. Tobin. A variable dimension solution approach for the general spatial equilibrium problem. *Mathematical Programming*, 40:33–51, 1988.

[37] H. R. Varian. *Microeconomic Analysis*. W.W. Norton & Company, New York, New York, 1978.