Mathematical Programs with Complementarity Constraints

Sven Leyffer*

Jan. 7, 2003

1 Introduction

An exciting new application of nonlinear programming techniques, are Mathematical Programs with Complementarity Constraints (MPCC),

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & c(x) \geq 0 \\ & 0 \leq x_1 \ \perp \ x_2 \geq 0, \end{array} \tag{1.1}$$

where $x = (x_0, x_1, x_2)$ and $\perp$ is the complementarity operator which requires that either a component $x_{1i} = 0$ or the corresponding component $x_{2i} = 0$. It is straightforward to include equality constraints in (1.1). Problems of this type arise in many engineering and economic applications, see the survey [7], the monographs [13, 14] and the growing collections of test problems [10, 5].

One attractive way of solving (1.1) is to replace the complementarity condition by a set of nonlinear inequalities, such as $X_1 x_2 \leq 0$ and then solve the equivalent Nonlinear Program (NLP),

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & c(x) \geq 0 \\ & x_1, \ x_2 \geq 0, \ X_1 x_2 \leq 0, \end{array} \tag{1.2}$$

where $X_1 = \text{diag}(x_1)$. Unfortunately, it can be shown [4, 16] that (1.2) violates the Mangasarian Fromovitz Constraint Qualification (MFCQ) at *any* feasible point. This failure of MFCQ implies that the multiplier set is unbounded, the central path fails to exist, the active constraint normals are linearly dependent and that linearizations of (1.2) can become inconsistent *arbitrarily close* to a solution. In addition, early numerical experience with this approach has been disappointing [2]. As a consequence, solving MPCCs via NLPs such as (1.2) has been commonly regarded as numerically unsafe.

The apparent failure of (1.2) has led to a number of special purpose algorithms for MPCCs, such as branch-and-bound [2], implicit nonsmooth approaches [14], piecewise SQP methods [13] and perturbation and penalization approaches [6] analyzed in [17]. However, all of these techniques require significantly more work than a standard NLP approach to (1.2)

Recently, there have been exciting new developments that have demonstrated that the gloomy prognosis against the use of (1.2) may have been premature. Standard NLP solvers have been used to solve a large class of MPCCs, written as NLPs, reliably and efficiently. This short note surveys these novel developments and provides a summary of open questions and possible extensions of these ideas.

The remainder is organized as follows. The next section provides a summary of certain stationarity concepts for MPCCs and establishes an important relationship with the KKT conditions of (1.2). This relationship is pivotal in the success of NLP solvers. The development of two important classes of solvers, Sequential Quadratic Programming (SQP) and Interior Point Methods (IPMs) is charted in the subsequent two sections. The note concludes by providing a brief outlook and describing some open problems.

2 The NLP revolution starts

The resurgence of interest in the analysis of NLP solvers applied to (1.2) is motivated by the success of SQP methods in particular. A very simple but key observation due to Scholtes is that strong stationarity is equivalent to the KKT conditions of (1.2).

A point x^* is called *strongly stationary*, if and

*Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439, USA `leyffer@mcs.anl.gov`

only if there exist multipliers $\lambda \geq 0$, $\hat{\nu}_1$ and $\hat{\nu}_2$ such that

$$
\begin{aligned}
\nabla f^* - \nabla c^{*^T} \lambda - \begin{pmatrix} 0 \\ \hat{\nu}_1 \\ \hat{\nu}_2 \end{pmatrix} &= 0 \\
c(x^*) &\geq 0 \\
x_1^*, x_2^* &\geq 0 \\
x_{1j}^* = 0 \;\text{ or }\; x_{2j}^* &= 0 \\
c_i^* \lambda_i = 0,\; x_{1j}^* \hat{\nu}_{1j} = 0,\; x_{2j}^* \hat{\nu}_{2j} &= 0 \\
\hat{\nu}_{1j} \geq 0,\; \hat{\nu}_{2j} \geq 0,\; \text{if } x_{1j}^* = x_{2j}^* &= 0.
\end{aligned} \tag{2.1}
$$

These are the KKT conditions of the *relaxed NLP* [16], which contains no complementarity condition and is therefor well-behaved.

The KKT conditions of (1.2), are very similar to (2.1) and this similarity will be exploited. Formally, a point x^* is called is a KKT point of (1.2), if and only if there exist multipliers $\lambda \geq 0$, $\nu_1 \geq 0$, $\nu_2 \geq 0$ and $\xi \geq 0$, such that

$$
\begin{aligned}
\nabla f^* - \nabla c^{*^T} \lambda - \begin{pmatrix} 0 \\ \nu_1 - X_2 \xi \\ \nu_2 - X_1 \xi \end{pmatrix} &= 0 \\
c(x^*) &\geq 0 \\
x_1^*, x_2^* &\geq 0 \\
X_1^* x_2 &\leq 0 \\
c_i(x)\lambda_i = 0,\; x_{1j}\nu_{1j} = 0,\; x_{2j}\nu_{2j} &= 0.
\end{aligned} \tag{2.2}
$$

Note, that complementarity between ξ and $X_1^* x_2 \leq 0$ follows trivially. Now observe, that (2.1) and (2.2) are equivalent, if we set

$$
\hat{\nu}_1 = \nu_1 - X_2 \xi \tag{2.3}
$$

$$
\hat{\nu}_2 = \nu_2 - X_1 \xi. \tag{2.4}
$$

This shows that there exists a *minimal value* of ξ, namely

$$
\xi_i = \begin{cases} 0 & \text{if } x_{1i}^* = x_{2i}^* = 0 \\ \max\left(0, \dfrac{-\hat{\nu}_{1i}}{x_{2i}^*}\right) & \text{if } x_{2i}^* > 0 \\ \\ \max\left(0, \dfrac{-\hat{\nu}_{2i}}{x_{1i}^*}\right) & \text{if } x_{1i}^* > 0. \end{cases} \tag{2.5}
$$

From which it follows that the unboundedness of the multipliers of (1.2) has a very special structure: the multipliers form a ray.

The fact that the KKT conditions of (1.2) are equivalent to strong stationarity imply the existence of bounded multipliers. This can be exploited in the analysis of SQP methods and in the design of robust IPM methods for MPECs.

3 SQP lead the way

SQP methods have recently been shown to solve MPCCs reliably as NLPs, despite the common folklore that this approach is doomed. Over 150 problems were solved and the SQP solver obtained *quadratic* convergence for all but two problems [8].

This success of SQP methods has motivated renewed interest in the theoretical properties of SQP methods. In [1] it is shown that an SQP method with *elastic mode* converges locally. The key idea is to consider a penalized version of (1.2). The penalty problem satisfies MFCQ and near a strongly stationary point, a sufficiently large penalty parameter can be found, similar to (2.5). Convergence can thus be established using standard techniques.

In [9] it is shown that SQP converges superlinearly near a strongly stationary point. The proof is divided into two parts. First, it is shown that if $x_1^{(k)^T} x_2^{(k)} = 0$ at some iteration k, then it follows that the SQP approximation of (1.2) about this point is equivalent to the SQP approximation of the relaxed NLP. Since the latter is a well-behaved problem, superlinear convergence follows. The second part of the proof assumes that $x_1^{(k)^T} x_2^{(k)} > 0$ and it is shown that each QP basis remains bounded away from singularity and convergence follows using standard techniques.

One undesirable assumption in [9] is the assumption that all QP approximations are consistent. This is trivially true, if $x_1^{(k)^T} x_2^{(k)} = 0$ for some k and it can be shown to hold, if the lower level problem satisfies a certain mixed-P property [13]. In general, a simple heuristic is implemented which relaxes the linearization of the complementarity constraint and this has been used successfully in [8].

4 Interior Point Methods

In contrast to SQP methods, Interior Point Methods (IPM) are not as robust at solving MPCCs. Using default settings they solve about 80 % of MPCCs. This is still remarkable, considering that the constraint gradients are dependent and the central path fails to exist!

The reason for the non-existence of the central path is the complementarity constraint. Clearly,

$$x_1 \geq 0 \ , \ x_2 \geq 0 \ \text{ and } \ X_1 x_2 \leq 0$$

has no interior. As a consequence, multipliers can become unbounded, resulting in slow progress (if any) toward the solution.

Three approaches to remedy this are currently being investigated. The first two are related to [17] and either relax the complementarity constraint or penalize it. The third approach mixes a simple active set heuristic with the IPM to identify and remove indices of bi-active constraints ($x_{1i} = 0 = x_{2i}$) [3]. However, it is not clear at present, what convergence properties this approach possesses.

The relaxation scheme [12, 15] introduces a parameter $\tau > 0$ and relaxes the complementarity constraint to

$$x_1 \geq 0 \ , \ x_2 \geq 0 \ \text{ and } \ X_1 x_2 \leq \tau.$$

A standard primal-dual method is then applied and the parameter τ is controlled in conjunction with the barrier parameter. It can be shown that near a strongly stationary point, the multipliers remain bounded and the central path exists.

An alternative to relaxation is to introduce an ℓ_1 penalty for the complementarity constraint and add $\rho x_1^T x_2$ to the objective. The resulting penalized NLP satisfies MFCQ and is well behaved. In addition, since $x_1, x_2 \geq 0$, no absolute values are required and the problem is smooth. Near a strongly stationary point, a sufficiently large (but finite) penalty parameter exists and any IPM method converges to this stationary point. We are currently investigating techniques for updating the penalty parameter, ρ.

The two schemes can be shown to be equivalent in the sense that for every relaxation τ there exists a penalty parameter ρ such that both approaches give the same solution. However, we prefer to control the penalty parameter, as it allows us to control the multipliers directly.

Interior Point Methods for MPCCs have also been considered in [13], namely PIPA. This is a hybrid SQP-IPM method, which aims to remain interior *only* with respect to the variables in the complementarity constraint, by perturbing it to

$$x_1 \geq 0 \ , \ x_2 \geq 0 \ \text{ and } \ X_1 x_2 = \tau.$$

Note, the last constraint is an *equation*. It is possible to construct a simple example, where $x_1^* = x_2^* = 0$ and the central path fails to exist! Thus, this perturbation is only suitable for problems without bi-active constraints. In [11] another example is constructed which shows that PIPA may fail to converge, even when strict complementarity ($x_1^* + x_2^* > 0$) holds. The reason for this adverse behavior is the trust-region used in PIPA which is controlled by the norm of the infeasibility.

5 Conclusion & Future

The underlying theme of the previous two sections has been to show that *small modifications* enable NLP solvers to work for MPCCs. Both SQP methods and IPM solvers either perturb or penalize the complementarity constraint. The key to proving convergence in both cases is the equivalence between strong stationarity (2.1) and the KKT conditions of the NLP formulation (2.2).

The robust solution of MPCCs as NLPs has harnessed the power of large scale NLP solvers to this new and exciting class of problems. Despite this success, however, there still remain some open questions.

An important open question is whether global convergence results can be established and — more importantly — whether these results can be strengthened to provide convergence to B-stationary points [16]. For instance, it is easy to construct examples for which the NLP approaches converge to a feasible C-stationary point. Unfortunately, C-stationary points allow trivial first order descent directions (and are really a misnomer!). Convergence to B-stationary points which are not strongly stationary can be observed in practice, even though the multiplier of the complementarity constraint ξ diverges to infinity.

Finally, some MPCCs require global solutions to be obtained. For instance, in the context of brittle fracture identification, the global minimum corresponds to the first structural failure. Local minima are physically meaningless in this case. Finding global minima for large NLPs is a very challenging problem and success is likely to involve the use of robust NLP techniques in conjunction with complementarity solvers.

References

[1] M. Anitescu. On solving mathematical programs with complementarity constraints as nonlinear programs. Preprint ANL/MCS-P864-1200, MCS Division, Argonne National Laboratory, Argonne, IL, USA, 2000.

[2] J.F. Bard. Convex two-level optimization. *Mathematical Programming*, 40(1):15–27, 1988.

[3] H. Benson, D.F. Shanno, and R.V.D. Vanderbei. Interior-point methods for nonconvex nonlinear programming: Complementarity constraints. Technical Report ORFE 02-02, Princeton University, Operations Research and Financial Engineering, 2002.

[4] Y. Chen and M. Florian. The nonlinear bilevel programming problem: Formulations, regularity and optimality conditions. *Optimization*, 32:193–209, 1995.

[5] S.P. Dirkse. MPEC world. webpage, GAMS Development Corp., `www.gamsworld.org/mpec/`, 2001.

[6] S.P. Dirkse, M.C. Ferris, and A. Meeraus. Mathematical programs with equilibrium constraints: Automatic reformulation and solution via constraint optimization. Technical Report NA-02/11, Oxford University Computing Laboratory, July 2002.

[7] M.C. Ferris and J.S. Pang. Engineering and economic applications of complementarity problems. *SIAM Review*, 39(4):669–713, 1997.

[8] R. Fletcher and S. Leyffer. Numerical experience with solving MPECs as NLPs. Numerical Analysis Report NA/210, Department of Mathematics, University of Dundee, Dundee, DD1 4HN, UK, 2002.

[9] R. Fletcher, S. Leyffer, D. Ralph, and S. Scholtes. Local convergence of SQP methods for mathematical programs with equilibrium constraints. Numerical Analysis Report NA/209, Department of Mathematics, University of Dundee, Dundee, DD1 4HN, UK, May 2002.

[10] S. Leyffer. MacMPEC: AMPL collection of MPECs. webpage, `www.mcs.anl.gov/~leyffer/MacMPEC/`, 2000.

[11] S. Leyffer. The penalty interior point method fails to converge for mathematical programs with equilibrium constraints. Numerical Analysis Report NA/208, Department of Mathematics, University of Dundee, February 2002.

[12] X. Liu and Sun J. Generalized stationary points and an interior point method for mathematical programs with equilibrium constraints. Preprint, Singapore, School of Business, 2002.

[13] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, 1996.

[14] J. Outrata, M. Kocvara, and J. Zowe. *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*. Kluwer Academic Publishers, Dordrecht, 1998.

[15] A. Raghunathan and L. T. Biegler. Mpec formulations and algorithms in process engineering. Technical report, CMU Chemical Engineering, 2002.

[16] H. Scheel and S. Scholtes. Mathematical program with complementarity constraints: Stationarity, optimality and sensitivity. *Mathematics of Operations Research*, 25:1–22, 2000.

[17] S. Scholtes. Convergence properties of regularization schemes for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 11(4):918–936, 2001.