On the Quadratic Convergence of the Singular Newton's Method (Essay)

R. A. Tapia
Y. Zhang

CRPC-TR92294
December 1992

Center for Research on Parallel Computation
Rice University
P.O. Box 1892
Houston, TX 77251-1892

On the Quadratic Convergence of the Singular Newton's Method [1] (Essay)

R.A. Tapia
Y. Zhang

December, 1992

TR92-43

[1] In *SIAG/OPT Views-and-News: A Forum for the SIAM Activity Group in Optimization*, Ed. Larry Nazareth, No. 1, Fall 1992, pp. 6-8

On the Quadratic Convergence of the Singular Newton's Method

Richard Tapia [1] and Yin Zhang [2]

The purpose of this essay is to describe a situation that we have found particularly exciting in our recent work in interior-point methods for linear programming. To our surprise, we have seen considerable theory developed concerning the superlinear convergence of singular Newton's methods. Hopefully, our comments will motivate further research in the general area of fast convergence for the singular Newton's method.

What is the general perception of Newton's method? Well, we all know that it is a most wonderful algorithm for approximating the zeros of a nonlinear system of equations. Its numerical and theoretical properties are known to us and we understand both its strengths and its weaknesses. Under well-known standard conditions concerning smoothness and nonsingularity it is not hard to establish local and fast convergence. For years we have embraced these conditions and argue that they are both reasonable and mild. Indeed, in some ill-defined but meaningful sense they must be necessary and sufficient for local and fast convergence. Experience has shown us that the semi-local properties of the method are actually quite good; in fact much better than the theory predicts. Convergence and fast convergence are not restricted to a very small neighborhood of the solution as many vendors of awkward hybrid methods would have us believe. This experience has also taught us that damping the Newton step, i.e., choosing steplength less than one, often improves the global behavior of Newton's method. However, not choosing steplength one locally may preclude fast convergence. The concern for these two aspects of Newton's method has led to the so-called backtracking strategies. In such a strategy one always considers steplength one before damping and implements damping in a manner which takes steplength one near the solution.

The literature is actually quite sparse when it comes to satisfying results concerning singular Newton's method in finite dimensional spaces. Some of

[1] Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77251-1892.

[2] Department of Mathematics and Statistics, University of Maryland, Baltimore County Campus, Baltimore, Maryland, 21228.

the results that do exist make the assumption that the rank deficiency of the Jacobian is extremely small, e.g. one. This lack of satisfying theory and our numerical experience reinforce the general belief that local fast convergence is not to be enjoyed by the singular Newton's method. However, our current message is that we have missed something by being excessively general in the problem class considered. We now support this statement by briefly describing what we consider to be a very convincing theory for the singular Newton's method in the application area of primal-dual interior-point methods for linear programming. It seems quite natural to believe that there is a more general theory which contains the linear programming application as a special case. Further study and understanding in the area of singular Newton's method is merited.

Today the interior-point methods of choice for linear programming all have the basic structure of the primal-dual method originally proposed by Kojima, Mizuno and Yoshise [2] based on earlier work of Megiddo [5]. These methods can be viewed as perturbed damped singular Newton's method applied to the first-order conditions for a particular standard form linear program. In the standard form linear program, the only inequality constraints are nonnegativity constraints on the variables. This formulation also has the nice feature that the duality gap at a feasible point is equal to the ℓ_1-norm of the nonlinear residual (first-order conditions). Hence the duality gap is an excellent merit function; it is nonnegative and zero only at a solution.

The term perturbed describes the situation that at each iteration the right-hand side of the Newton equation is modified to accommodate so-called adherence to the central path behavior. The interior-point aspect of the algorithm comes from the fact that at each iteration the new iterate is forced to stay positive, i.e. strictly satisfy the inequality constraints by staying in the interior of the nonnegative orthant. This is accomplished by starting with a strictly feasible initial iterate and damping the Newton step at each iteration. An interesting feature here is that there is no guarantee that steplenghth one will ever be allowed, even near the solution.

The singularity of the Newton's method comes from the fact that the Jacobian of the system in question is singular at the solution for degenerate linear programs. Moreover, most real-world problems are known to be degenerate. Hence in practice degeneracy is the rule and not the exception. Moreover, the rank deficiency in many problems is quite large.

It seems that Newton's method is starting out with three strikes against it; namely forced perturbation, forced damping, and singularity. All three are natural enemies of fast (superlinear) convergence. Our perturbed damped singular Newton's method has two algorithmic parameters that the designer or user is free to choose. The first choice we shall denote by t and the second by u. The parameter t is strictly between 0 and 1 and denotes the percentage by which one chooses to move toward the boundary of the positive orthant. Specifically, $t = 0$ implies no movement and $t = 1$ implies movement onto the boundary. It should be appreciated that the choice $t = 1$ does not imply steplength 1. The parameter u merely designates the perturbation to the right-hand side of the Newton system.

The earliest theoretical papers on this topic, Kojima, Mizuno and Yoshise [2], Monteiro and Adler [8], and Todd and Ye [10], for example, all established polynomial complexity for various choices of the algorithmic parameters. From a Newton's method point of view their form of polynomial complexity implies global linear convergence of the duality gap sequence to zero. At this juncture global linear convergence is the most that should be expected from a perturbed damped singular Newton's method. Today we know that the linear convergence obtained was actually quite poor and tended to be negatively correlated with the quality of the complexity bounds. Could it be that fast convergence and polynomiality actually work against each other? This inconsistency was further fueled by the work of Lustig, Marsten and Shanno [3]. They deviated from the algorithmic parameter choices which were known to give polynomiality and successfully constructed fast algorithms. However, it was not clear if their form of the perturbed damped singular Newton method possessed polynomiality, indeed, if it was globally convergent.

The issue of superlinear convergence was brought into the mainstream of activity in February of 1990 at the Asilomar meeting when Zhang, Tapia, and Dennis (see [15]) presented two theories for superlinear convergence of the increasingly popular primal-dual interior-point methods (perturbed damped singular Newton's methods). Their first theory assumed nondegeneracy (equivalently nonsingularity) and gave conditions which the algorithmic parameters t and u should satisfy in order to guarantee quadratic convergence of the duality gap sequence to zero. A main issue here was the demonstration that it was possible to choose t (percentage to boundary) in a manner which allowed the steplength to approach 1 sufficiently fast so that the quadratic convergence of Newton's method was not lost. Their second theory did not use the assumption of nondegeneracy and gave conditions which the algorithmic parameters t and u should satisfy in order to guarantee superlinear convergence of the duality gap sequence to zero. Immediately some questioned the consistency of the Zhang-Tapia-Dennis assumptions. Others conjectured that polynomial complexity and superlinear convergence were incompatible. However, soon after Zhang and Tapia [14] squelched these doubts by constructing a class of choices for the algorithmic parameters and showing that for these choices the perturbed damped singular Newton method possessed polynomial complexity and gave superlinear convergence for degenerate problems. They showed that for nondegenerate problems polynomial complexity and quadratic convergence could be obtained. This was particularly satisfying since no one really expected quadratic convergence for degenerate problems. However, there was one annoying aspect to this situation — in practice quadratic convergence was observed even for degenerate problems. Hence our story continues.

Mizuno, Todd and Ye [7] considered a variant of the Newton method we have been discussing and established superior polynomial complexity bounds for this variant. They called this variant a predictor-corrector method. The predictor-correct aspect of the algorithm entailed two linear solves per iteration. Hence when comparing convergence rate results for the Mizuno-Todd-Ye predictor-corrector method with those for the standard method, they should technically be considered as two-step results. The predictor-corrector variant can also be viewed as a perturbed damped singular Newton's method.

Ye, Tapia and Zhang [13] showed that the Mizuno-Todd-Ye predictor-corrector algorithm gave superlinear convergence for degenerate problems and quadratic convergence for nondegenerate problems while maintaining its superior complexity. McShane [4] independently derived a similar result. Ye, Güler, Tapia and Zhang [12], based on Ye, Tapia, Zhang [13], were able to demonstrate the surprising result that the Mizuno-Todd-Ye predictor-corrector algorithm actually gave quadratic convergence in all cases including degenerate problems. Mehrotra [6], also based on Ye, Tapia and Zhang [13], independently established a similar result. So technically we now have two-step quadratic convergence for a perturbed damped singular Newton's method. This result motivated several very strong and intense attempts to establish an analogous one-step result.

Zhang and Tapia [16] extended the Zhang-Tapia-Dennis theory in [15] and dropped the assumption of nondegeneracy. They were able to use this new theory to construct a perturbed damped Newton method which demonstrably had an order of convergence r for any $1 \leq r < 2$. While their theory gave conditions for quadratic convergence they were not able to construct such an algorithm. Ye [11], working with the basic model of the Mizuno-Todd-Ye predictor-corrector method, was able to construct one-step versions with convergence order r for any r satisfying $1 \leq r \leq 2$. However, his algorithm with a convergence order of 2 is only subquadratic since the Q_2-factor is actually infinite.

The final chapter to this exciting story is presently being written. Very recently Gonzaga and Tapia [1] working with the Mizuno-Todd-Ye predictor-corrector primal-dual interior-point method constructed a variant which can be viewed as a perturbed damped singular Newton method and gives (one-step) quadratic convergence.

We hope that the exciting and intense activity that led to superlinear convergence theory for various forms of singular Newton's methods in linear programming will spread to more general problem areas.

References

1. C. Gonzaga and R.A. Tapia. A quadratically convergent primal-dual algorithm for linear programming. Technical Report, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, in preparation.

2. M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior-point for linear programming. In Nimrod Megiddo, editor, Progress in Mathematical programming, interior-point and related methods, pages 29-47. Springer-Verlag, New York, 1989.

3. I.J. Lustig, R.E. Marsten, and D.F. Shanno. Computational experience with a primal-dual interior-point method for linear programming. Linear Algebra and Its Applications, 152:191-222, 1991.

4. A. McShane. A superlinearly convergent $O(\sqrt{n}L)$ iteration primal-dual linear programming algorithm. Manuscript, 2537 Villanova Drive, Vienna, Va. 1991.

5. N. Megiddo. Pathways to the optimal set in linear programming. In Nimrod Megiddo, editor, Progress in Mathematical programming,

interior-point and related methods, pages 131-158. Springer-Verlag, New York, 1989.

6. S. Mehrotra. Quadratic convergence in a primal-dual method. Technical Report 91-15, Dept. of Industrial Engineering and Management Science, Northwestern University, 1991.

7. S. Mizuno, M.J. Todd and Y. Ye. On adaptive-step primal-dual interior-point algorithms for linear programming. Technical Report No. 944, School of ORIE, Cornell University, Ithaca, New York, 1990.

8. R.C. Monteiro and I. Adler. Interior path-following primal-dual algorithms. Part I: linear programming. Math. Prog., 44:27-41, 1989.

9. R.A. Tapia, Y. Zhang, and Y. Ye. On the convergence of the iteration sequence in primal-dual interior-point methods. Technical Report No. 91-24, Dept. of Computational and Applied Mathematics, Rice University, 1991.

10. M.J. Todd and Y. Ye. A centered projective algorithm for linear programming. Math. of O.R., 15:508-529, 1990.

11. Y. Ye. On the Q-order convergence of interior-point algorithms for linear programming. Dept. of Management Sciences, The University of Iowa, 1991.

12. Y. Ye, O. Güler, R.A. Tapia, and Y. Zhang. A quadratically convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming. Technical Report TR91-26, Dept. Computational and Applied Mathematics, Rice University, 1991.

13. Y. Ye, R.A. Tapia, and Y. Zhang. A superlinearly convergent $O(\sqrt{n}L)$-iteration algorithm for linear programming. Technical Report TR91-22, Dept. Computational and Applied Mathematics, Rice University, 1991.

14. Y. Zhang and R.A. Tapia. A superlinearly convergent polynomial primal-dual interior-point algorithm for linear programming. Technical Report TR90-40, Dept. Computational and Applied Mathematics, Rice University, 1990.

15. Y. Zhang, R.A. Tapia, and J.E. Dennis. On the superlinear and quadratic convergence of primal-dual interior point linear programming algorithms. SIAM Journal on Optimization, 2:303-324, 1992.

16. Y. Zhang and R.A. Tapia. Superlinear and quadratic convergence of primal-dual interior-point methods for linear programming revisited. Journal of Optimization: Theory and Applications, 73:229-242, 1992.