Referee's report:
Adaptive barrier strategies for nonlinear interior methods
Jorge Nocedal, Andreas Wächter and Richard A. Waltz

The authors present strategies for adaptively adjusting the barrier parameter in interior-point methods for nonlinear programming, which in my opinion is a very interesting line of research. The investigation is done by an extension of predictor-corrector methods for linear programming. The reduction of μ is measured by progress with respect to a partially linearized approximation of the residual of the nonlinear equations that firm the first-order optimality conditions. The authors embed their scheme in a "classical" reduction scheme to ensure convergence. They also include a discussion on the suitability of the corrector step for linear programming, when the initial point is not particularly well chosen.

The convergence strategy resembles *spacer steps* [Lue84], which is not a new idea. The idea is to use a known method to guarantee convergence (monotone mode) and make sure that the proposed method (free mode) does not inhibit convergence. A feature here is that the free steps do not contribute at all to the convergence. We could replace the l monotone steps by l steps of (almost) any kind. I do not find this very appealing. What is there to say that the approach that the authors suggest is particularly good? For example, would the minimizing σ in (4.2) be a particularly good choice?

I do not fully understand the convergence argument after the globalization framework. The authors claim that *"... it is clear that $\Phi_k \to 0$"*. As I state below in item 4, I find the use of k misleading. It is not totally clear what work is involved in each step. Is it true with no assumptions that all limit points satisfy the first-order optimality conditions, as stated in the following sentence? I would prefer a more specific globalization discussion.

Specific comments

1. It is mentioned that [7,20] provide certain convergence results to stationary points. Isn't that exactly what is done here too? The merit function is the residual of the first-order optimality conditions.

2. The authors talk about solving (3.3). Yes, but what if no solution exist, or if the inertia is wrong? This is explained much later.

3. Page 5, line (3.5)+4.
I do not understand what "conversely" refers to. Is it σ small implies good progress?

4. I find the use of k confusing in the table of the globalization framework. I would think of k being one iteration, i.e., one set of (x, y, z). I assume that this is what is tabulated in the figures. However, for the monotone mode, you may perform many iterations before the convergence criteria is satisfied. Hence, the update of k by one appears very odd to me. (As it is now, I would solve the problem with $k \leq 2$ if κ is sufficiently small.)

5. It appears to me that there is a minus sign missing in (7.1).

6. I do not follow the discussion in the paragraph following (7.3). It is the right-hand side of (3.3), and not (2.5), as far as I can see, since $\mu = 0$ for the affine-scaling step.

7. It is not clear to me that the corrector step is applied at the primal-dual point. Rather, the corrector step and the affine scaling step are both independent of μ. The corrector step is aimed at improving the affine scaling step. To me, they form one entity aimed at optimality, and the step incurred by μ is aimed at centrality. It is not clear to me that the corrector step of Figure 4 is unfavorable. If we add the affine-scaling step and the corrector step, we are very close to the optimum. Ideally, they would give the optimum together, so I do not see why this step is unfavorable.

8. I do not understand the results of Table 1. For LP, we would expect a dual variable to diverge if the corresponding primal inequality constrained has the satisfied with equality on the feasible region, and symmetrically for the primal variables. I am not very familiar with the NETLIB test set, but this may be the case for FORPLAN. However, I would not expect the objective function values to diverge. I am not totally sure what disabling the initial point procedure in PCx means. The initial point of ones does not appear very bad to me, so I do not understand why PCx would fail for this choice of initial point.

9. If predictor-corrector methods for linear programs are so sensitive for the choice of starting point, is this the case also for nonlinear programs? I cannot see described how you go about defining initial strictly positive x and z. I assume that you may have to add slack variables,

or something similar, if the constraint is not strictly positive, and then the initial value might be crucial.

References

[Lue84] D. G. Luenberger. *Linear and Nonlinear Programming.* Addison-Wesley Publishing Company, Reading, second edition, 1984. ISBN 0-201-15794-2.