ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois 60439

# The Return of the Active Set Method

**Sven Leyffer**

Mathematics and Computer Science Division

Preprint ANL/MCS-P1277-0805

February 17, 2005, revised August 5, 2005

[1]This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contract W-31-109-ENG-38.

# The Return of the Active Set Method*

Sven Leyffer†

February 17, 2005, revised August 5, 2005

For solving nonlinear optimization problems, two competing iterative approaches are available: active set methods and interior-point methods. Current implementations of interior methods often outperform active set methods in terms of speed. On the other hand, active set methods are more robust and better suited for warm starts, which are important for solving integer optimization problems [8, 9]. Consequently, we have recently become interested in new active set approaches, which are reviewed in this note.

## 1 Active Set Methods for Quadratic Programs

Consider the quadratic programming (QP) problem

$$\underset{x}{\text{minimize}} \ \frac{1}{2}x^T Hx + g^T x \quad \text{subject to} \ \ A^T x = b \ \text{ and } \ l \leq x \leq u,$$

where $A \in \mathbb{R}^{n \times m}$ has full rank and $H$ is symmetric but not necessarily positive definite. Our new active set approach has two main components.

First, we identify an estimate of the optimal active set by approximately minimizing the augmented Lagrangian

$$L(x,y,\rho) := \frac{1}{2}x^T Hx + g^T x - y^T(A^T x - b) + \frac{1}{2}\rho \|A^T x - b\|^2,$$

in the box $l \leq x \leq u$, where $y$ are the multipliers of $A^T x = b$ and $\rho_k$ is the penalty parameter. This step provides a Cauchy-point $x_c^k$, a first-order multiplier estimate $y_c^k = y^k - \rho_k(A^T x_c^k - b)$, and an active set estimate $\mathcal{A}_k := \{i : [x_c^k]_i = l_i \ \text{ or } [x_c^k]_i = u_i\}$. This step is similar to the iterates generated by LANCELOT [3].

Next, we solve an equality constraint QP in the remaining inactive variables indexed by $\mathcal{I} := \{1,\ldots,n\} \ \mathcal{A}$ by computing an approximate solution to the first-order conditions

$$\begin{bmatrix} H_{\mathcal{I},\mathcal{I}} & -A_{:,\mathcal{I}} \\ A^T_{:,\mathcal{I}} & \end{bmatrix} \begin{pmatrix} \Delta x_{\mathcal{I}} \\ \Delta y \end{pmatrix} = - \begin{pmatrix} [\nabla x L(x_k^c, y_k^c, 0)]_{\mathcal{I}} \\ A^T x_k^c - b \end{pmatrix},$$

where $H_{\mathcal{I},\mathcal{I}}$ is the submatrix of $H$ corresponding to rows and columns of $\mathcal{I}$. We then perform a backtracking line-search along $(x_c^k + \alpha \Delta x_{\mathcal{I}}, y_c^k + \alpha \Delta y)$ to ensure global convergence. We show that if $\alpha = 1$, then the two steps are equivalent to a Newton step on the first-order conditions.

*Preprint ANL/MCS-P1277-0805; to appear in Oberwolfach Reports, 2005.

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

Global convergence is enforced through the use of a filter [6, 7]. A filter $\mathcal{F}$ is a list of pairs of constraint violation $h^l := \|A^T x^l - b\|$ and first order error $\theta^l := \| \min \left( \nabla_x L(x,y) \ , \ x \right) \|$, where $L(x,y)$ is the Lagrangian of the QP. A new point $x^k$ is acceptable to the filter if $h^k \leq \beta h^l$, or $\theta^k \leq \beta \theta^l$ for all $l \in \mathcal{F}$. The backtracking line-search reduces $\alpha$ until an acceptable iterate is found. This acceptable iterate may be added to the filter.

Traditional proofs for augmented Lagrangian methods use two forcing sequences $\eta_k \searrow 0$ and $\omega_k \searrow 0$ to control progress in $h^k$ and $\theta^k$, respectively, and guide the penalty parameter. Recently, it has been shown [4] that $\eta_k \searrow 0$ is sufficient to ensure convergence for certain QPs. The filter approach removes the need for any forcing sequence whose choice may be problematic in practice. Preliminary numerical experience is encouraging, and we are able to detect the optimal active set in a modest number of iterations.

## 2 Active Set Methods for Nonlinear Programs

Recently, researchers have expressed renewed interest in sequential linear programming (SLP) methods for nonlinear optimization problems such as

$$\underset{x}{\text{minimize}} \ \ f(x) \quad \text{subject to} \ \ c(x) \geq 0;$$

see [5, 2, 1]. These SLP methods solve a trust-region LP around the current iterate $x^k$, given by

$$\underset{d}{\text{minimize}} \ \ g_k^T d \quad \text{subject to} \ \ c_k + A_k^T d \geq 0 \ \ \text{and} \ \ \|d\|_\infty \leq \Delta_k,$$

where $g_k = \nabla f(x^k)$, $c_k = c(x^k)$, and $A_k = \nabla c(x^k)^T$. The solution of this LP provides an estimate of the active inequality constraints, which is used to define an equality constrained QP to compute a second-order step.

One problematic aspect of this approach is the use of the $\ell_\infty$ trust-region. It has been observed that while the active constraints corresponding to $c(x) \geq 0$ settle down, the active trust-region bounds do not, and this feature may cause the LP solver to perform many wasteful pivots even close to the solution.

We propose an alternative trust-region subproblem based on penalizing an elliptic or $\ell_2$ trust-region. This gives rise to the following active set identification problem

$$\underset{d}{\text{minimize}} \ \ \mu g_k^T d + \frac{1}{2} d^T d \quad \text{subject to} \ \ c_k + A_k^T d \geq 0.$$

It can be shown that the dual of this problem is a bound-constrained QP in the multipliers $y$,

$$\underset{y}{\text{minimize}} \ \ \frac{1}{2} y^T A^T A y - (c - \mu A^T g)^T y + \frac{\mu^2}{2} g^T g \quad \text{subject to} \ \ y \geq 0.$$

Convergence of a filter algorithm along the lines of [2] can be shown. The proof exploits a piecewise quadratic relationship between the penalty parameter $\mu$ and the $\ell_2$ trust-region radius $\Delta$.

## 3 Conclusions

We have introduced two active set identification strategies for optimization. Both schemes can be implemented in a matrix-free format, requiring only matrix-vector operations and iterative linear system solves. We believe that this is an important ingredient for a successful large-scale active set strategy.

## References


[1] R. H. Byrd, N. I. M. Gould, J. Nocedal, and R. A. Waltz. An algorithm for nonlinear optimization using linear programming and equality constrained subproblems. *Mathematical Programming, Series B*, 100(1):27–48, 2004.

[2] C.M. Chin and R. Fletcher. On the global convergence of an SLP-filter algorithm that takes EQP steps. *Mathematical Programming*, 96(1):161–177, 2003.

[3] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. *Trust-Region Methods.* MPS-SIAM Series on Optimization. SIAM, Philadelphia, 2000.

[4] Z. Dostal, A. Friedlander, and S. A. Santos. Augmented Lagrangian with adaptive precision control for quadratic programming with equality constraints. *Computational Optimization and Applications*, 14(1):37–53, 1999.

[5] R. Fletcher and E. Sainz de la Maza. Nonlinear programming and nonsmooth optimization by successive linear programming. *Math. Programming*, 43:235–256, 1989.

[6] R. Fletcher and S. Leyffer. Nonlinear programming without a penalty function. *Math. Programming*, 91:239–270, 2002.

[7] R. Fletcher, S. Leyffer, and Ph. L. Toint. On the global convergence of a filter-SQP algorithm. *SIAM J. Opt.*, 13(1):44–59, 2002.

[8] C. A. Floudas. *Nonlinear and Mixed–Integer Optimization.* Topics in Chemical Engineering. Oxford University Press, New York, 1995.

[9] I. E. Grossmann. Review of nonlinear mixed-integer and disjunctive programming techniques. *Optimization and Engineering*, 3:227–252, 2002.