QPBLUR: A Regularized Active-Set Method for Sparse Convex Quadratic Programming

Christopher Maes and Michael Saunders

Gurobi Optimization

SOL and ICME, Stanford University

ISMP 2012

21st International Symposium on Mathematical Programming

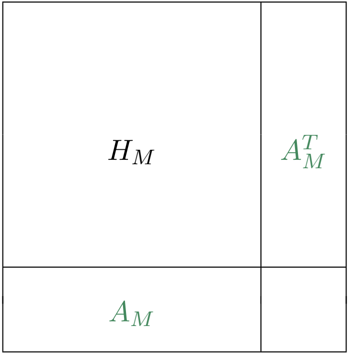
TUB, Berlin, August 19–24, 2012

1. QPBLUR overview
2. Active-set methods for quadratic programming
3. Regularization
4. Bound-Constrained Lagrangian method (BCL)
5. Computational results
6. Contributions and future work

QPBLUR overview

Convex Quadratic Program ($H \in \mathrm{R}^{n \times n}, A \in \mathrm{R}^{m \times n}, H \succeq 0$)

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^T x + \tfrac{1}{2} x^T H x \\ \text{subject to} \quad & Ax = b, \quad \ell \le x \le u \end{aligned}$$



A_M = cols of A corresponding to Moving variables

QPBLUR overview

Convex Quadratic Program $(H \in \mathrm{R}^{n\times n}, A \in \mathrm{R}^{m\times n}, H \succeq 0)$

$$\begin{array}{ll} \underset{x,\,y}{\text{minimize}} & c^T x + \frac{1}{2} x^T H x + \frac{\delta}{2} \|x\|_2^2 + \frac{\mu}{2} \|y\|_2^2 \\ \text{subject to} & Ax + \mu y = b, \quad \ell \le x \le u \end{array}$$

$$\begin{bmatrix} H_M + \delta I & A_M^T \\ A_M & -\mu I \end{bmatrix}$$

Regularization makes KKT system nonsingular

We use augmented Lagrangian to remove μ, y

References

Regularization for LP, QP

PDQ1 (S 1996), OSL (S, Tomlin 1996), PDCO (S 1998 . . .)
HOPDM (Altman, Gondzio 1999)
Friedlander and Orban 2012
Greif, Moulding and Orban 2012

Augmented Lagrangian methods

LANCELOT (Conn, Gould, Toint 1992)
Optimal QP algorithms (Dostál 2009)

QP bibliography and forthcoming book

ftp://ftp.numerical.rl.ac.uk/pub/qpbook/qpbook.bib
Computational Quadratic Programming (Gould, Toint 2012)

References

Hanh Huynh (QPBLU)
A Large-scale Quadratic Programming Solver Based on Block-LU Updates of the KKT system
SCCM, Stanford University, Sep 2008

Christopher Maes (QPBLUR)
A Regularized Active-set Method for Sparse Convex Quadratic Programming
ICME, Stanford University, Nov 2010

Elizabeth Wong (icQP)
Active-Set Methods for Quadratic Programming
Dept of Mathematics, UC San Diego, Jun 2011

Active-set methods for quadratic programming

Nullspace method (SQOPT)

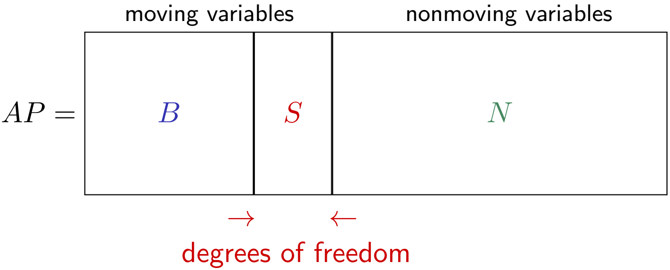
SQOPT: a nullspace (reduced-gradient) method

$$\text{minimize} \quad c^T x + \tfrac{1}{2} x^T H x \quad \text{subject to} \quad Ax = b, \quad \ell \le x \le u$$

Nullspace method (SQOPT)

SQOPT: a nullspace (reduced-gradient) method

$$\text{minimize} \quad c^T x + \tfrac{1}{2} x^T H x \quad \text{subject to} \quad Ax = b, \quad \ell \le x \le u$$



Limitations of SQOPT

All active-set methods solve for a search direction in the moving variables M

$$\begin{pmatrix} H_M & A_M^T \\ A_M & \end{pmatrix} \begin{pmatrix} \Delta x_M \\ \Delta y \end{pmatrix} \approx \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

Limitations of SQOPT

All active-set methods solve for a search direction in the moving variables M

$$\begin{pmatrix} H_M & A_M^T \\ A_M & \end{pmatrix} \begin{pmatrix} \Delta x_M \\ \Delta y \end{pmatrix} \approx \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

Reduced-gradient method uses a specific ordering (Gill et al. 1990):

$$A_M = \begin{pmatrix} B & S \end{pmatrix} \quad \to \quad \begin{pmatrix} H_{BB} & B^T & H_{BS} \\ B & & S \\ H_{SB} & S^T & H_{SS} \end{pmatrix}$$

Reduced Hessian is the Schur complement of blue block:

$$Z_M^T H_M Z_M = H_{SS} - (\dots)(\dots)^{-1}(\dots)^T$$

Limitations of SQOPT

All active-set methods solve for a search direction in the moving variables M

$$\begin{pmatrix} H_M & A_M^T \\ A_M & \end{pmatrix} \begin{pmatrix} \Delta x_M \\ \Delta y \end{pmatrix} \approx \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

Reduced-gradient method uses a specific ordering (Gill et al. 1990):

$$A_M = \begin{pmatrix} B & S \end{pmatrix} \quad \rightarrow \quad \begin{pmatrix} H_{BB} & B^T & H_{BS} \\ B & & S \\ H_{SB} & S^T & H_{SS} \end{pmatrix}$$

Reduced Hessian is the Schur complement of blue block:

$$Z_M^T H_M Z_M = H_{SS} - (\dots)(\dots)^{-1}(\dots)^T$$

Best when degrees of freedom small (say < 2000)

Likely to be dense $\Rightarrow$ Need to work with original KKT system

Fullspace method (QPBLU)

QPBLU uses block-LU updates to the KKT factors (Huynh 2008)

$$K_0 = \begin{pmatrix} H_0 & A_0^T \\ A_0 & \end{pmatrix} = L_0 D_0 {L_0}^T \quad \text{or} \quad L_0 U_0$$

Fullspace method (QPBLU)

QPBLU uses block-LU updates to the KKT factors (Huynh 2008)

$$K_0 = \begin{pmatrix} H_0 & A_0^T \\ A_0 & \end{pmatrix} = L_0 D_0 L_0{}^T \quad \text{or} \quad L_0 U_0$$

To change moving variables, work with bordered system:

$$\begin{pmatrix} K_0 & V \\ V^T & D \end{pmatrix} = \begin{pmatrix} L_0 & \\ Z^T & I \end{pmatrix} \begin{pmatrix} U_0 & Y \\ & C \end{pmatrix}$$

Fullspace method (QPBLU)

QPBLU uses block-LU updates to the KKT factors (Huynh 2008)

$$K_0 = \begin{pmatrix} H_0 & A_0^T \\ A_0 & \end{pmatrix} = L_0 D_0 {L_0}^T \quad \text{or} \quad L_0 U_0$$

To change moving variables, work with bordered system:

$$\begin{pmatrix} K_0 & V \\ V^T & D \end{pmatrix} = \begin{pmatrix} L_0 & \\ Z^T & I \end{pmatrix} \begin{pmatrix} U_0 & Y \\ & C \end{pmatrix}$$

- Inertia-controlling method of (Gill 2007) keeps K_0 nonsingular
- Quasi-Newton updates to H_0 handled same way
- Y, Z sparse, C dense but small ($\leq$ no. of updates to K_0)
- L_0, U_0 from a sparse direct solver

Limitations of QPBLU

- Theoretically, inertia-controlling method ensures K_0 nonsingular
- Singular (or nearly singular) K_0 is a difficulty in practice
- Must detect singularity (need rank-revealing factorization)
- LUSOL, MA48, MA27, MA57 have threshold rook pivoting but tight threshold $\Rightarrow$ dense factors
- KKT repair complicated (and costly)

Regularization

Primal and dual regularization

For convex QP, small $\delta, \mu > 0 \Rightarrow$ nonsingular KKT systems:

$$\begin{pmatrix} H_M + \delta I & A_M^T \\ A_M & -\mu I \end{pmatrix}$$

- Primal regularization

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \phi(x) + \tfrac{1}{2}\delta \left\| x \right\|_2^2 \\ \text{subject to} & Ax = b, \quad \ell \le x \le u \end{array}$$

  Under certain conditions, optimal x unaltered for $\delta \le \bar{\delta}$

  (Friedlander and Tseng 2007)

- Dual regularization

$$\begin{array}{ll} \underset{x,y}{\text{minimize}} & \phi(x) + \tfrac{1}{2}\mu \left\| y \right\|_2^2 \\ \text{subject to} & Ax + \mu y = b, \quad \ell \le x \le u \end{array}$$

Regularization in interior methods for LP and QP

- Fixed regularization

  PDQ1 (S 1996), OSL (S & Tomlin 1996), PDCO (S 2002 . . .)

$$\underset{x,\,y}{\text{minimize}} \qquad c^T x + \tfrac{1}{2}\gamma \left\|x\right\|^2 + \tfrac{1}{2}\delta \left\|y\right\|^2$$

$$Ax + \delta y = b, \qquad \ell \le x \le u$$

Indefinite LDL^T on KKT is stable if γ, δ sufficiently large

Regularization in interior methods for LP and QP

- Fixed regularization

  PDQ1 (S 1996), OSL (S & Tomlin 1996), PDCO (S 2002 . . .)

$$\underset{x,\,y}{\text{minimize}} \quad c^T x + \tfrac{1}{2}\gamma \|x\|^2 + \tfrac{1}{2}\delta \|y\|^2$$

$$Ax + \delta y = b, \qquad \ell \le x \le u$$

Indefinite LDL^T on KKT is stable if γ, δ sufficiently large

- Dynamic regularization

  HOPDM (Altman and Gondzio 1999)

Smaller perturbation via *dynamic proximal point* terms:

$$\tfrac{1}{2}(x - x_k)^T R_p (x - x_k), \qquad \tfrac{1}{2}(y - y_k)^T R_d (y - x_y)$$

QPBLUR version 1

Sequence of regularized QPs with $\mu_k, \delta_k \to 0$

$$\begin{array}{ll} \underset{x,y}{\text{minimize}} & c^T x + \tfrac{1}{2} x^T H x + \tfrac{1}{2}\delta_k \|x\|_2^2 + \tfrac{1}{2}\mu_k \|y\|_2^2 \\ \text{subject to} & Ax + \mu_k y = b, \qquad \ell \le x \le u \end{array}$$

KKT systems

$$\begin{pmatrix} H_M + \delta_k I & A_M^T \\ A_M & -\mu_k I \end{pmatrix} \begin{pmatrix} \Delta x_M \\ -\Delta y \end{pmatrix} = \begin{pmatrix} -g_M + A_M^T y \\ 0 \end{pmatrix}$$

- QPs strictly convex
- Always feasible (no Phase 1)
- KKTs nonsingular for any set M (any A_M)
- Use any sparse LU or LDL$^\text{T}$ solver (black-box)
- Use QPBLU's block-LU updates without change

Bound-Constrained Lagrangian method (BCL)

LANCELOT

A BCL method for constrained NLP (Conn, Gould, Toint 1992)

$$\text{minimize } \varphi(x) \quad \text{subject to} \quad c(x) = 0, \quad \ell \le x \le u$$

Augmented Lagrangian (Hestenes 1969, Powell 1969, Bertsekas 1982)

$$\mathcal{L}(x, \hat{y}, \rho) = \varphi(x) - \hat{y}^T c(x) + \tfrac{1}{2}\rho \, \|c(x)\|_2^2$$

LANCELOT

A BCL method for constrained NLP (Conn, Gould, Toint 1992)

$$\text{minimize } \varphi(x) \quad \text{subject to} \quad c(x) = 0, \quad \ell \le x \le u$$

Augmented Lagrangian (Hestenes 1969, Powell 1969, Bertsekas 1982)

$$\mathcal{L}(x, \hat{y}, \rho) = \varphi(x) - \hat{y}^T c(x) + \tfrac{1}{2}\rho \, \|c(x)\|_2^2$$

- Subproblem: minimize $\mathcal{L}(x, \hat{y}, \rho)$ st $\ell \le x \le u$
- Solve to get $\hat{x}$ (optimality tol $\omega \to 0$)
- If $\|c(\hat{x})\| < \eta$, update $\hat{y}$ (feasibility tol $\eta \to 0$)
  else increase ρ
- Repeat

LANCELOT

A BCL method for constrained NLP (Conn, Gould, Toint 1992)

$$\text{minimize } \varphi(x) \quad \text{subject to} \quad c(x) = 0, \quad \ell \le x \le u$$

Augmented Lagrangian (Hestenes 1969, Powell 1969, Bertsekas 1982)

$$\mathcal{L}(x, \hat{y}, \rho) = \varphi(x) - \hat{y}^T c(x) + \tfrac{1}{2}\rho \, \|c(x)\|_2^2$$

- Subproblem: minimize $\mathcal{L}(x, \hat{y}, \rho)$ st $\ell \le x \le u$
- Solve to get $\hat{x}$ (optimality tol $\omega \to 0$)
- If $\|c(\hat{x})\| < \eta$, update $\hat{y}$ (feasibility tol $\eta \to 0$)
  else increase ρ
- Repeat

BCL method keeps $\rho < \infty$

QPBLUR version 2: Regularized BCL for QP

Sequence of regularized BCL subproblems (δ, $\hat{y}$, ρ changing)

$$\begin{array}{ll} \underset{x,\,r}{\text{minimize}} & c^T x + \tfrac{1}{2} x^T H x + \tfrac{1}{2}\delta \|x\|_2^2 + \hat{y}^T r + \tfrac{1}{2}\rho \|r\|_2^2 \\ \text{subject to} & Ax + r = b, \quad \ell \le x \le u \end{array}$$

QPBLUR version 2: Regularized BCL for QP

Sequence of regularized BCL subproblems (δ, $\hat{y}$, ρ changing)

$$\begin{array}{ll} \underset{x,\,r}{\text{minimize}} & c^T x + \tfrac{1}{2} x^T H x + \tfrac{1}{2}\delta \|x\|_2^2 + \hat{y}^T r + \tfrac{1}{2}\rho \|r\|_2^2 \\ \text{subject to} & Ax + r = b, \quad \ell \le x \le u \end{array}$$

Regularized BCL method

- Solve subproblem to get $\hat{x}, \hat{r}, \hat{z}$ (optimality tol $\omega \to \omega^\star$)
- If $\|\hat{r}\| < \eta$, update $\hat{y} \leftarrow \hat{y} + \rho\hat{r}$ (feasibility tol $\eta \to \eta^\star$)
  else increase ρ
- Stop if $\|\hat{r}\| \le \eta^\star$ and $\left\|\nabla\phi(\hat{x}) - A^T\hat{y} - \hat{z}\right\| \le \omega^\star$
- Repeat ($\delta \to \delta^\star$)

Active-set method for BCL subproblem

Larger KKT system

$$\begin{pmatrix} H_M + \delta I & & A_M^T \\ & \rho I & I \\ A_M & I & \end{pmatrix} \begin{pmatrix} \Delta x_M \\ \Delta r \\ -\Delta y \end{pmatrix} = \begin{pmatrix} -g_M + A_M^T y \\ -\hat{y} - \rho r + y \\ 0 \end{pmatrix}$$

Eliminate Δr

$$\begin{pmatrix} H_M + \delta I & A_M^T \\ A_M & -\frac{1}{\rho} I \end{pmatrix} \begin{pmatrix} \Delta x_M \\ -\Delta y \end{pmatrix} = \begin{pmatrix} -g_M + A_M^T y \\ \frac{1}{\rho}(\hat{y} - y) + r \end{pmatrix}$$

- Same KKT matrix as before ($1/\rho \equiv \mu$), but different rhs
- Start from any active set (any submatrix A_M)
- Any sparse LU or LDL^T linear solver
- Block-LU updates without change

QPBLUR software

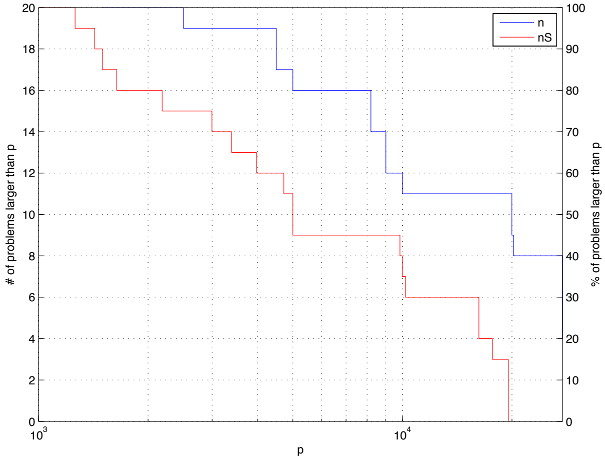
- MATLAB and Fortran 95 versions
- `bclsolve:` Outer BCL algorithm
- `regularqp:` Regularized BCL subproblem solver
- May be used as bound-constrained LS solver
  (avoids forming $A^T A$; a sparse version of LSSOL)
- Block-LU updates and sparse linear solver interface (Hunyh 2008)

  LUSOL, MA57, PARDISO (parallel), SuperLU, UMFPACK

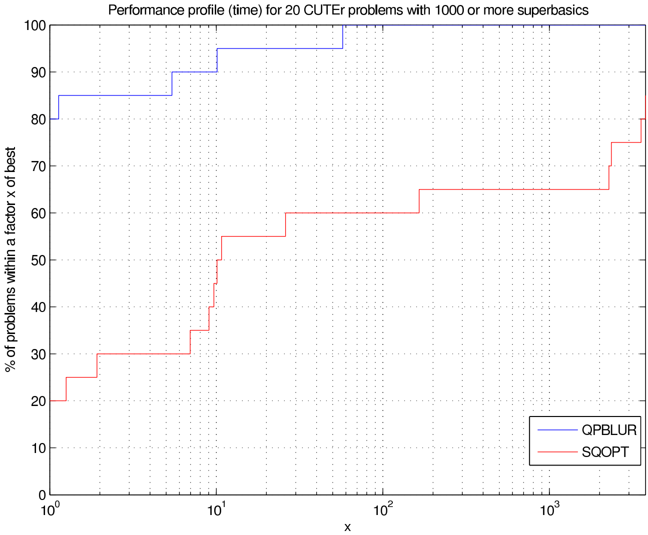
Computational results

CUTEr (Constraint and Unconstrained Test Environment revisited)

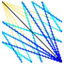
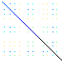
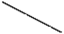
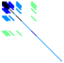
- 20 convex QPs with 1000 or more degrees of freedom
- Compare QPBLUR to SQOPT



QPs with many degrees of freedom



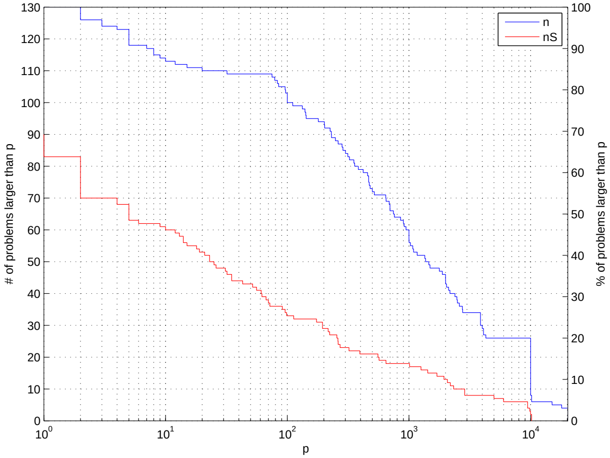
QPs with many degrees of freedom

Problem	A	H	m	n	dof	SQOPT	QPBLUR
AUG2DC			10000	20200	10200	2024(10622)	0.5(3)
AUG2DCQP			10000	20200	9994	2283(14296)	236(10441)
AUG3DCQP			10000	20200	17677	3537(22071)	391(11554)
CVXQP2			10000	20200	2186	127(8770)	692(7714)
DTOC3			2998	4499	1499	9.2(1830)	0.1(2)
MOSARQP2			700	2500	1640	6.2(2610)	3.2(878)
STCQP2			4095	8193	3970	180(8080)	6.9(132)

CPU seconds (# of itns)

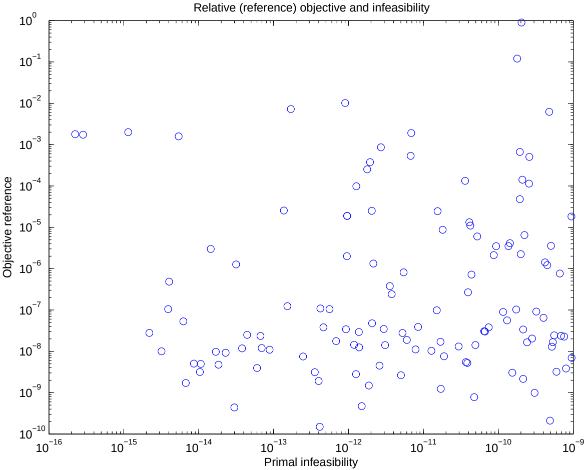
Maros and Mészáros convex QPs

- Test set of 138 convex quadratic programs (1999)
- Created to test interior-point QP solver BPMPD
- Includes reference optimal objective value
- Most problems have only a few degrees of freedom



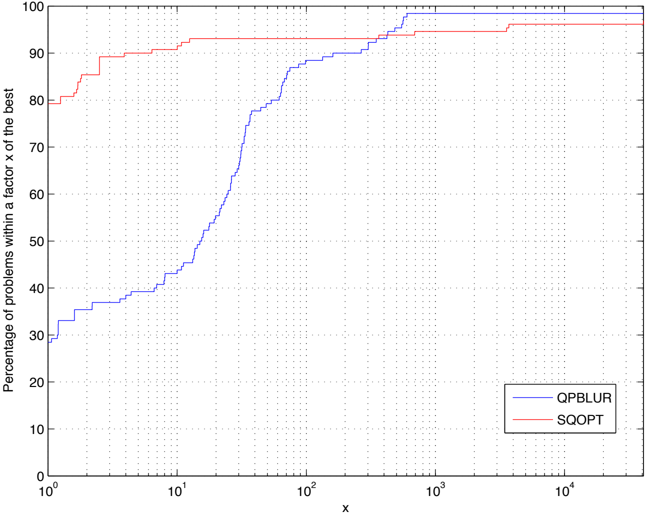
Maros and Mészáros convex QPs

- Excluded 8 problems (CUTEr errors, disastrous fill-in)
- QPBLUR solves 128/130 problems
- Plot of (ϵ_r, ϕ_r): $\epsilon_r = \frac{\|Ax-s\|_\infty}{1+\|(x,s)\|_\infty}$ $\quad \phi_r = \frac{\left|\phi_B^\star - \phi_Q^\star\right|}{1+\left|\phi_B^\star\right|}$



Maros and Mészáros convex QPs

- QPBLUR solves 128/130 problems
- QPBLUR versus SQOPT (time)



Comparison with LANCELOT on QPs

LANCELOT solves the NLP

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \varphi(x) \\ \text{subject to} & c(x) = 0, \quad \ell \le x \le u \end{array}$$

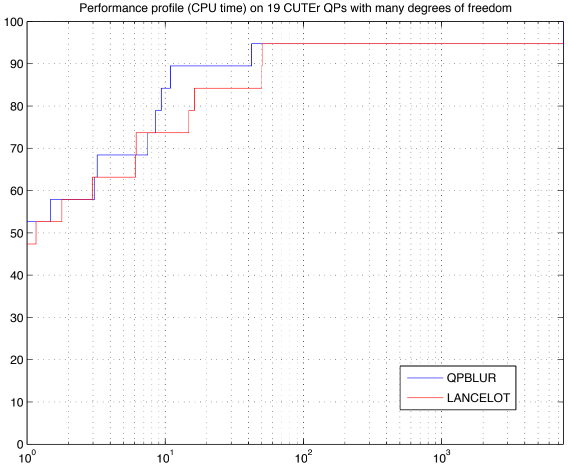
by solving BCL subproblems

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \varphi(x) - y_k^T c(x) + \frac{1}{2}\rho_k \|c(x)\|_2^2 \\ \text{subject to} & \ell \le x \le u \end{array}$$

- LANCELOT uses SBMIN subproblem solver:
  trust-region with gradient-projection method
  CG with band preconditioner (MA57 optional)
- QPBLUR uses the same outer algorithm (BCL)
  adjusting primal regularization each BCL iteration

Comparison with LANCELOT on QPs

- 20 convex CUTEr QPs with many degrees of freedom
- LANCELOT option `quadratic-problem YES`
- QPBLUR versus LANCELOT (time)



QPBLUR inside SNOPT

SNOPT's SQP method for NLPs (using first derivatives)

$$\underset{x}{\text{minimize}}\ \varphi(x) \quad \text{subject to} \quad \ell \le \begin{pmatrix} x \\ c(x) \\ Ax \end{pmatrix} \le u$$

QPBLUR inside SNOPT

SNOPT's SQP method for NLPs (using first derivatives)

$$\underset{x}{\text{minimize}}\ \varphi(x) \quad \text{subject to} \quad \ell \le \begin{pmatrix} x \\ c(x) \\ Ax \end{pmatrix} \le u$$

QP subproblems

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \varphi_k + g_k^T(x - x_k) + \tfrac{1}{2}(x - x_k)^T H_k (x - x_k) \\ \text{subject to} \quad & \text{linearized constraints} + \text{bounds} \end{aligned}$$

QPBLUR inside SNOPT

SNOPT's SQP method for NLPs (using first derivatives)

$$\underset{x}{\text{minimize}}\ \varphi(x) \quad \text{subject to} \quad \ell \le \begin{pmatrix} x \\ c(x) \\ Ax \end{pmatrix} \le u$$

QP subproblems

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \varphi_k + g_k^T(x - x_k) + \frac{1}{2}(x - x_k)^T H_k (x - x_k) \\ \text{subject to} & \text{linearized constraints} + \text{bounds} \end{array}$$

- Warm start QPBLUR with (x, y) and active set from previous QP

QPBLUR inside SNOPT

SNOPT's SQP method for NLPs (using first derivatives)

$$\underset{x}{\text{minimize}}\ \varphi(x) \quad \text{subject to} \quad \ell \le \begin{pmatrix} x \\ c(x) \\ Ax \end{pmatrix} \le u$$

QP subproblems

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \varphi_k + g_k^T(x - x_k) + \frac{1}{2}(x - x_k)^T H_k (x - x_k) \\ \text{subject to} & \text{linearized constraints} + \text{bounds} \end{array}$$

- Warm start QPBLUR with (x, y) and active set from previous QP
- Block-LU KKT updates handle quasi-Newton $H_k = G_k^T G_k$ as well as active-set changes

QPBLUR inside SNOPT

SNOPT's SQP method for NLPs (using first derivatives)

$$\underset{x}{\text{minimize}}\ \varphi(x) \quad \text{subject to} \quad \ell \leq \begin{pmatrix} x \\ c(x) \\ Ax \end{pmatrix} \leq u$$

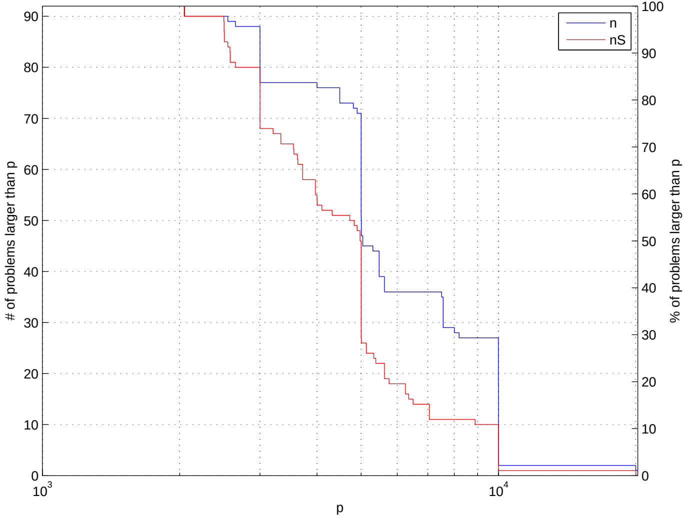
QP subproblems

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \varphi_k + g_k^T(x - x_k) + \frac{1}{2}(x - x_k)^T H_k (x - x_k) \\ \text{subject to} & \text{linearized constraints} + \text{bounds} \end{array}$$

- Warm start QPBLUR with (x, y) and active set from previous QP
- Block-LU KKT updates handle quasi-Newton $H_k = G_k^T G_k$ as well as active-set changes
- If QP infeasible, revert to SQOPT (elastic bounds on linearized constraints)

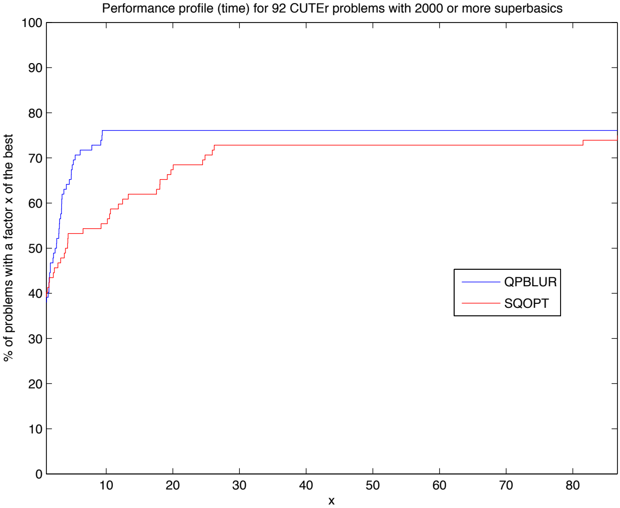
Nonlinear CUTEr problems

- 92 problems with 2000 or more degrees of freedom $(m \ll n)$



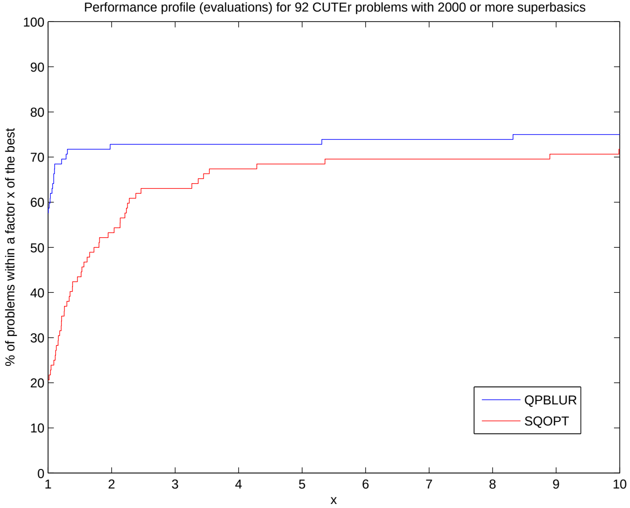
Nonlinear CUTEr problems

- QPBLUR using UMFPACK
- SNOPT-QPBLUR versus SNOPT-SQOPT (time)



Nonlinear CUTEr problems

- Functions $\varphi(x)$, $c(x)$, $\nabla\varphi, \nabla c_i(x)$ can be expensive
- SNOPT-QPBLUR versus SNOPT-SQOPT (functions)



JANNSON3 problem

$$\begin{aligned}
\text{minimize} \quad & \tfrac{1}{2}x_1^4 + \tfrac{1}{2}x_2^2 + (1 - x_1^2) + \sum_{i=1}^{2n}(x_i - 1)^2 \\
& + \underline{d}_1 x_1 + \underline{d}_2 x_2 - \underline{x}_1^2 \underline{x}_2 - \underline{d}_1 \underline{x}_1 - \underline{d}_2 \underline{x}_2 \\
\text{subject to} \quad & \sum_{i=1}^{2n} x_i - 1 = 0, \quad \sum_{i=1}^{2n} x_1^2 - 0.75 \le 0 \\
& (\underline{x}^2)^T x^1 + (\underline{x}^1)^T x^2 - (\underline{x}^1)^T \underline{x}^2 - 1/(5n) \le 0 \\
& (\overline{x}^2)^T x^1 + (\overline{x}^1)^T x^2 - (\overline{x}^1)^T \underline{x}^2 - 1/(5n) \le 0 \\
& -1 \le x_i \le 1, \quad i = 1, \ldots, 2n
\end{aligned}$$

20000 variables, 3 nonlinear constraints, 1 linear constraint + bounds
19999 degrees of freedom at solution

SNOPT-QPBLUR solved the problem in

8	major iterations
2 or 3	BCL iterations per major
1	QPBLUR iteration per BCL subproblem
10	seconds

Contributions and future work

Contributions

- Active-set method + KKT systems allows many degrees of freedom
- Block-LU updates allow any black-box LU (preferably with separate L and U)
- Regularization ensures nonsingular KKT systems
- BCL algorithm (sequence of regularized QP subproblems) moderates the perturbation to the original problem

Contributions

- Active-set method + KKT systems allows many degrees of freedom
- Block-LU updates allow any black-box LU (preferably with separate L and U)
- Regularization ensures nonsingular KKT systems
- BCL algorithm (sequence of regularized QP subproblems) moderates the perturbation to the original problem

QPBLUR takes full advantage of convexity to permit a simplified algorithm (for each QP subproblem)

- No need for Phase 1
- No need to control inertia (no KKT repair)
- No danger of cycling in the presence of degeneracy (trivial step-length procedure, bounds satisfied exactly)
- Can warm start from any point and any active set

Future work (SNOPT-QPBLUR integration)

- Handle elastic mode without reverting to SQOPT

$$\begin{array}{ll} \underset{x,\,r}{\text{minimize}} & \varphi(x) + \gamma \, \|r\|_1 \\ \text{subject to} & \ell \le \begin{pmatrix} x \\ c(x) + r \\ Ax \end{pmatrix} \le u \end{array}$$

$$\begin{array}{ll} \underset{x,\,s,\,r}{\text{minimize}} & \varphi_k + g_k^T(x - x_k) + \frac{1}{2}(x - x_k)^T H_k (x - x_k) + \gamma \, \|r\|_1 \\ \text{subject to} & c_k + J_k(x - x_k) - s + r = 0 \\ & \ell \le \begin{pmatrix} x \\ s \\ Ax \end{pmatrix} \le u \end{array}$$

Future work (SNOPT-QPBLUR integration)

- Handle elastic mode without reverting to SQOPT

$$\begin{array}{ll} \underset{x,r}{\text{minimize}} & \varphi(x) + \gamma \|r\|_2^2 \\ \text{subject to} & \ell \le \begin{pmatrix} x \\ c(x) + r \\ Ax \end{pmatrix} \le u \end{array}$$

$$\begin{array}{ll} \underset{x,s,r}{\text{minimize}} & \varphi_k + g_k^T(x - x_k) + \frac{1}{2}(x - x_k)^T H_k (x - x_k) + \gamma \|r\|_2^2 \\ \text{subject to} & c_k + J_k(x - x_k) - s + r = 0 \\ & \ell \le \begin{pmatrix} x \\ s \\ Ax \end{pmatrix} \le u \end{array}$$

Future work (SNOPT-QPBLUR integration)

- Handle elastic mode without reverting to SQOPT

$$\begin{array}{ll} \underset{x,\,r}{\text{minimize}} & \varphi(x) + \gamma \left\| r \right\|_2^2 \\ \text{subject to} & \ell \le \begin{pmatrix} x \\ c(x) + r \\ Ax \end{pmatrix} \le u \end{array}$$

$$\begin{array}{ll} \underset{x,\,s,\,r}{\text{minimize}} & \varphi_k + g_k^T(x - x_k) + \frac{1}{2}(x - x_k)^T H_k (x - x_k) + \gamma \left\| r \right\|_2^2 \\ \text{subject to} & c_k + J_k(x - x_k) - s + r = 0 \\ & \ell \le \begin{pmatrix} x \\ s \\ Ax \end{pmatrix} \le u \end{array}$$

- Use SQOPT when dof small, switch to QPBLUR (or icQP!)

Future work (SNOPT-QPBLUR integration)

- Handle elastic mode without reverting to SQOPT

$$\begin{array}{ll} \underset{x,\,r}{\text{minimize}} & \varphi(x) + \delta \left\| x \right\|_2^2 \\ \text{subject to} & \ell \le \begin{pmatrix} x \\ c(x) + r \\ Ax \end{pmatrix} \le u \end{array}$$

$$\begin{array}{ll} \underset{x,\,s,\,r}{\text{minimize}} & \varphi_k + g_k^T(x - x_k) + \frac{1}{2}(x - x_k)^T H_k (x - x_k) + \delta \left\| x \right\|_2^2 \\ \text{subject to} & c_k + J_k(x - x_k) - s + r = 0 \\ & \ell \le \begin{pmatrix} x \\ s \\ Ax \end{pmatrix} \le u \end{array}$$

- Use SQOPT when dof small, switch to QPBLUR (or icQP!)
- Make SNOPT aware of primal regularization δ

Future work (QPBLUR)

- Relax stopping criterion based on dual optimality condition of original QP?
- Detailed sparse solver performance study: LUSOL, MA57, PARDISO, SuperLU, UMFPACK, MUMPS
- Block-LDL^T updates (take advantage of symmetry)
- Develop gradient projection subproblem solver (allow for multiple changes to the active-set)

References

Hanh Huynh (QPBLU)
A Large-scale Quadratic Programming Solver Based on Block-LU Updates of the KKT system
SCCM, Stanford University, Sep 2008

Christopher Maes (QPBLUR)
A Regularized Active-set Method for Sparse Convex Quadratic Programming
ICME, Stanford University, Nov 2010

Elizabeth Wong (icQP)
Active-Set Methods for Quadratic Programming
Dept of Mathematics, UC San Diego, Jun 2011

References

Hanh Huynh (QPBLU)
A Large-scale Quadratic Programming Solver Based on Block-LU Updates of the KKT system
SCCM, Stanford University, Sep 2008

Christopher Maes (QPBLUR)
A Regularized Active-set Method for Sparse Convex Quadratic Programming
ICME, Stanford University, Nov 2010

Elizabeth Wong (icQP)
Active-Set Methods for Quadratic Programming
Dept of Mathematics, UC San Diego, Jun 2011

Philip Gill and Elizabeth Wong

- SQIC = SQOPT + icQP
- Reduced-gradient method or Block-LU updates of KKT
- Convex and nonconvex QP
- Aiming to handle 2nd derivatives in SNOPT

References

Hanh Huynh (QPBLU)
A Large-scale Quadratic Programming Solver Based on Block-LU Updates of the KKT system
SCCM, Stanford University, Sep 2008

Christopher Maes (QPBLUR)
A Regularized Active-set Method for Sparse Convex Quadratic Programming
ICME, Stanford University, Nov 2010

Elizabeth Wong (icQP)
Active-Set Methods for Quadratic Programming
Dept of Mathematics, UC San Diego, Jun 2011

Philip Gill and Elizabeth Wong

- SQIC = SQOPT + icQP
- Reduced-gradient method or Block-LU updates of KKT
- Convex and nonconvex QP
- Aiming to handle 2nd derivatives in SNOPT

One day: SQICr = SQOPT + icQP + QPBLUR??

References

Hanh Huynh (QPBLU)
A Large-scale Quadratic Programming Solver Based on Block-LU Updates of the KKT system
SCCM, Stanford University, Sep 2008

Christopher Maes (QPBLUR)
A Regularized Active-set Method for Sparse Convex Quadratic Programming
ICME, Stanford University, Nov 2010

Elizabeth Wong (icQP)
Active-Set Methods for Quadratic Programming
Dept of Mathematics, UC San Diego, Jun 2011

Philip Gill and Elizabeth Wong

- SQIC = SQOPT + icQP
- Reduced-gradient method or Block-LU updates of KKT
- Convex and nonconvex QP
- Aiming to handle 2nd derivatives in SNOPT

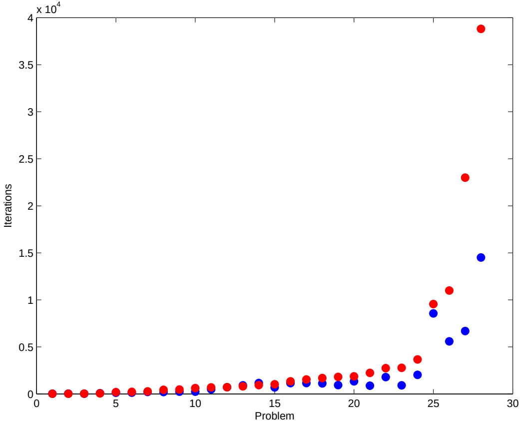
One day: SQICr = SQOPT + icQP + QPBLUR??

Funding: COMSOL, ONR, ARL, DOE

Extra slides

28 Infeasible problems

- Infeasible LP test problems (Netlib `lpi_xxx`)
- Results from April 2009 (Banff saddle point systems meeting)



$\bullet \rho \nearrow 10^{12}$ $\quad \bullet \rho \equiv 1$

Quickly changing the active-set

Typical active-set methods add/delete a constraint per iteration.

- k_0 constraints active at x_0 and $k^\star$ constraints active at $x^\star$
- Require at least $|k_0 - k^\star|$ iterations

Gradient projection on bound constrained problems

- $x_{k+1} =$ project $x_k - \alpha \nabla \phi(x_k)$ onto $\{x : \ell \leq x \leq u\}$
- Define active-set via constraints active at x_{k+1}
- Optimize in subspace of free variables

Allows for large changes to the active-set

Apply Gradient projection to BCL subproblem: $\underset{\ell \leq x \leq u}{\text{minimize}} \ \mathcal{L}(x, \hat{y}, \rho)$

- Solve same regularized KKT system to accurately optimize in subspace
- If active-set changes dramatically refactor KKT; otherwise use BLU updates

Regularized QP subproblem

Bound-constrained subproblem

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^T x + \frac{1}{2} x^T H_\delta x + \hat{y}^T (b - Ax) + \frac{1}{2} \left\| {D_\rho}^{1/2} (b - Ax) \right\|_2^2 \\ \text{subject to} & \ell \leq x \leq u \end{array}$$

with $H_\delta \succ 0$, $D_\rho \succ 0$ and diagonal

Regularized QP subproblem

Bound-constrained subproblem

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^T x + \tfrac{1}{2} x^T H_\delta x + \hat{y}^T (b - Ax) + \tfrac{1}{2} \left\| {D_\rho}^{1/2} (b - Ax) \right\|_2^2 \\ \text{subject to} \quad & \ell \le x \le u \end{aligned}$$

with $H_\delta \succ 0$, $D_\rho \succ 0$ and diagonal

$\equiv$ Linearly constrained subproblem

$$\begin{aligned} \underset{x,\,r}{\text{minimize}} \quad & c^T x + \tfrac{1}{2} x^T H_\delta x + \hat{y}^T r + \tfrac{1}{2} r^T D_\rho r \\ \text{subject to} \quad & Ax + r = b, \qquad \ell \le x \le u \end{aligned}$$

- Prefer to work with constrained problem for numerical reasons
- Can always satisfy $Ax + r = b$, $\ell \le x \le u$

Active-set method for QP subproblem

Yields larger KKT system

$$\begin{pmatrix} H_{\delta M} & & A_M^T \\ & D_\rho & I \\ A_M & I & \end{pmatrix} \begin{pmatrix} \Delta x_M \\ \Delta r \\ -\Delta y \end{pmatrix} = \begin{pmatrix} -g_M + A_M^T y \\ -\hat{y} - D_\rho r + y \\ 0 \end{pmatrix}$$

Eliminate Δr and we have

$$\begin{pmatrix} H_{\delta M} & A_M^T \\ A_M & -D_\rho^{-1} \end{pmatrix} \begin{pmatrix} \Delta x_M \\ -\Delta y \end{pmatrix} = \begin{pmatrix} -g_M + A_M^T y \\ D_\rho^{-1}(\hat{y} - y) + r \end{pmatrix}$$

- KKT systems always nonsingular
- Can start from any set M (any matrix A_M)
- Can use any sparse LU or $\mathrm{LDL^T}$ factorization

Solver for sparse least squares problems

- Method for sparse least squares problems that avoids forming $H = A^T A$
- Sparse version of LSSOL

Solver for sparse least squares problems

- Method for sparse least squares problems
  that avoids forming $H = A^T A$
- Sparse version of LSSOL

Tikhonov regularized bound-constrained least squares

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \|b - Ax\|_2^2 + \gamma \|x\|_2^2 \\ \text{subject to} & \ell \le x \le u \end{array}$$

- Method for sparse least squares problems that avoids forming $H = A^T A$
- Sparse version of LSSOL

Tikhonov regularized bound-constrained least squares

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \|b - Ax\|_2^2 + \gamma \|x\|_2^2 \\ \text{subject to} & \ell \leq x \leq u \end{array}$$

$\equiv$ Regularized QP subproblem

$$\begin{array}{ll} \underset{x,r}{\text{minimize}} & \gamma \frac{1}{2} x^T x + \frac{1}{2} r^T r \\ \text{subject to} & Ax + r = b, \quad \ell \leq x \leq u \end{array}$$

Can solve weighted LS problems with covariance matrix $W \succ 0$