Preconditioner Updates for Solving Sequences of Indefinite Linear Systems in Optimization

Benedetta Morini
Università degli Studi di Firenze

Based on works with Stefania Bellavia, Valentina De Simone, Daniela di Serafino, Margherita Porcelli

Recent Advances on Optimization, CERFACS-Toulouse, July 24-26, 2013

Outline

Consider the problem of preconditioning a sequence of linear systems

$$\mathcal{A}_k x = b_k, \quad k = 1, \ldots$$

where $\mathcal{A}_k \in \mathbb{R}^{n\times n}$ are nonsingular indefinite sparse matrices.

- Computing preconditioners $\mathcal{P}_1, \mathcal{P}_2, \ldots$, for individual systems separately can be very expensive.

  A reduction of the cost can be achieved by sharing some of the computational effort among subsequent linear systems.

- Given an algebraic preconditioner $\mathcal{P}_{seed}$ for some *seed* matrix $\mathcal{A}_{seed}$ of the sequence, we investigate how to form updated preconditioners for subsequent matrices $\mathcal{A}_k$ at a low computational cost.

Outline (c.ed)

Updating strategies are an alternative to freezing the preconditioner.

A periodical or dynamic refresh of the seed preconditioner may be necessary.

Content of the talk:

1. State of art in preconditioning update techniques for nonsymmetric and symmetric linear systems.
2. New proposals for updating preconditioners for two classes of systems:
   - nonsymmetric linear systems arising in Newton-Krylov methods;
   - KKT systems arising in Interior Point methods.

Updating frameworks in literature

Limited-memory Quasi-Newton preconditioners:

- symmetric positive definite (SPD) matrices and nonsymmetric matrices arising in Newton methods:
  [Morales, Nocedal 2000], [Bergamaschi, Bru, Martinez, Putti 2006], [Gratton, Sartenaer, Tshimanga 2011].

Recycled Krylov information preconditioners:

- symmetric and nonsymmetric matrices:
  [Carpentieri, Duff, Giraud 2003], [Knoll, Keyes, 2004], [Loghin, Ruiz, Tohuami 2006], [Giraud, Gratton, Martin, 2007], [Fasano, Roma 2013].

Incremental ILU preconditioners:

- nonsymmetric matrices: [Calgaro, Chehab, Saad 2010].

Updates of factorized preconditioners:

- SPD matrices and nonsymmetric matrices:
  [Meurant 2001], [Benzi, Bertaccini 2003], [Duintjer Tebbens, Tuma 2007, 2010], [Bellavia, Bertaccini, M. 2011], [Bellavia, De Simone, di Serafino, M. 2011-2012]

Approximate updates of factorized preconditioners

Consider two linear systems

$$\mathcal{A}_{seed} x = b, \qquad \mathcal{A}_k x = b_k$$

and let $\mathcal{P}_{seed} = LDU \approx \mathcal{A}_{seed}$.

- It follows

$$\mathcal{A}_k = \mathcal{A}_{seed} + (\mathcal{A}_k - \mathcal{A}_{seed}) \approx L(\,D + \underbrace{L^{-1}(\mathcal{A}_k - \mathcal{A}_{seed})U^{-1}}_{\text{ideal update}}\,)U$$

- The *ideal* update of the middle-term is costly:
  - the difference matrix $\mathcal{A}_k - \mathcal{A}_{seed}$ should be formed;
  - in general the ideal update is dense and its factorization is impractical.
- Form an *approximate* and cheap update.

Preconditioning & Matrix-free setting

- Unpreconditioned Krylov methods are matrix-free.

  But a truly matrix-free setting is lost when an algebraic preconditioner is used.

- A preconditioning strategy is classified as *nearly* matrix-free if it lies close to a true matrix-free settings. Specifically, if
  - only a few full matrices are formed;
  - for preconditioning most of the systems of the sequence, matrices that are reduced in complexity with respect to the full $\mathcal{A}_k's$ are required.

  [Knoll, Keyes 2004]

- Nearly matrix-free updating strategies have been proposed.

Update of LDU factorizations [Duintjer Tebbens, Tuma 2007, 2010]

Ideal updated preconditioner for $\mathcal{A}_k$:

$$\mathcal{A}_k \approx L(D + \underbrace{L^{-1}(\mathcal{A}_k - \mathcal{A}_{seed})U^{-1}})U$$

The approximate updated preconditioner is obtained as follows:

1. Neglect either L^{-1} or U^{-1} (closeness of L or U to the identity matrix):

$$\begin{aligned} \mathcal{A}_k &\approx L(\,D + (\mathcal{A}_k - \mathcal{A}_{seed})U^{-1}\,)U \\ \mathcal{A}_k &\approx L(\,D + L^{-1}(\mathcal{A}_k - \mathcal{A}_{seed})\,)U \end{aligned}$$

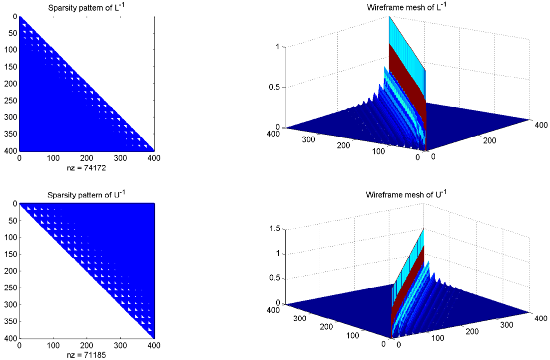
2. Use only a triangular part of the current matrix $\mathcal{A}_k$:

$$\begin{aligned} \mathcal{P}_k &= L(\,DU + triu(\mathcal{A}_k - \mathcal{A}_{seed})) \\ \mathcal{P}_k &= (L\,D + tril(\mathcal{A}_k - \mathcal{A}_{seed})\,)U \end{aligned}$$

$\mathcal{P}_k$ is factorized. This approach is not suitable for symmetric matrices.

Motivation: matrices with decaying inverses:

- banded SPD and indefinite matrices [Demko, Moss, Smith 1984][Meurant 1992];
- nonsymmetric block tridiagonal matrices [Nabben 1999];
- matrices $h(A)$ with A symm and banded and h analytic [Benzi, Golub 99].



2D Nonlinear Convection diffusion problem. Sparsity pattern (on the left) and wireframe mesh (on the right) of the inverses of the L and U factors obtained from the ILU factorization of the Jacobian at the null vector ($n = 400$)

Banded approximate factors

Ideal updated preconditioner for $\mathcal{A}_k$:

$$\mathcal{A}_k \approx L(D + \underbrace{L^{-1}(\mathcal{A}_k - \mathcal{A}_{seed})U^{-1}})U$$

The approximate updated preconditioner is obtained as follows:

1. Let $f(M) = \mathrm{band}(M, k_l, k_u)$, be the banded approximation of M with k_l lower and k_u upper diagonals.

2. Let

$$E_k = f(\mathcal{A}_k - \mathcal{A}_{seed}), \qquad F_k = f(L^{-1}\, E_k\, U^{-1}),$$

and

$$\mathcal{P}_k = L\,(D + F_k)\,U.$$

[Benzi, Golub 1999], [Benzi, Bertaccini 2003], [Bellavia, Bertaccini, M. 2011], [Bellavia, M., Porcelli 2013]

- Small bandwidth values k_l and k_u are viable.

  The computationally most convenient approximations E_k and F_k are diagonal ($k_l = k_u = 0$).

- Forming/approximating L^{-1} and U^{-1}:
  - Use the Approximate INVerse (AINV) preconditioner [Benzi, Meyer, Tuma 1996], [Benzi, Tuma 1998]

    $$\mathcal{P}_{seed} = W\, D^{-1}\, Z^T \approx \mathcal{A}_{seed}^{-1}$$

    The updated $\mathcal{P}_k$ takes the form

    $$\mathcal{P}_k = W\, (D + F_k)^{-1}\, Z^T.$$

    [Benzi, Bertaccini 2003], [Bertaccini 2004], [Bellavia, Bertaccini, M. 2011]

  - Use banded approximation of L^{-1} and U^{-1}, computable without the neeed of a complete inversion of L^{-1} and U^{-1}.
    [Bellavia, M., Porcelli 2013]

Sequences of symmetric systems

- KKT matrices: we are aware of low-rank updates of the factorizations of the blocks, [Griewank, Walther, Korzec 2007]
- Updating techniques for SPD matrices:
  - Update of the factorized preconditioner for the Schur complement in Interior Point (IP) methods for linear programming,
    [Baryamureeba, Steihaug, Zhang 1999], [Wang, O'Leary 2000]
  - Factorization preconditioner updates for diagonally modified matrices arising in regularizing optimization methods and bound-constrained convex optimization,
    [Meurant 2001], [Bellavia, De Simone, di Serafino, M. 2011, 2012]

We devise and analyze modifications of the existing approaches for sequences of SPD matrices which can be used for sequences of nonsymmetric and KKT matrices respectively.

Sequences of systems in Newton-Krylov methods

$$F(x) = 0$$

$F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ continuously differentiable, J Jacobian matrix of F.

Sequence of Newton equations

$$J(x_k)s = -F(x_k), \quad k = 0, 1, \dots$$

- By continuity, $\{J(x_k)\}$ varies slowly if the iterates are close enough.
- Generally, $J(x_k)$ is nonsymmetric.
- Let $\mathcal{P}_{seed} = LDU$.

  Discard the off-diagonals of $J_k - J_{seed}$ from the ideal update

$$J_k = J_{seed} + (J_k - J_{seed}) \simeq LDU + \underbrace{diag(J_k - J_{seed})}_{\Sigma_k = diag(\sigma^k_{11}, \dots, \sigma^k_{nn})}$$

Approximate diagonally modified sequences

$$C = LDL^T + \Sigma_k$$

LDL^T is symmetric positive definite
Σ_k is diagonal positive semidefinite.

1. Form an approximate factorization for C setting

$$\mathcal{P}_k = L_k D_k L_k^T$$

with

$$\begin{aligned} D_k &= D + \Sigma_k, \\ L_k &= eye(n), \quad off(L_k) = off(L) Z_k \\ Z_k &= diag(z_{11}^k, \ldots, z_{nn}^k), \quad z_{ii}^k = \frac{d_{ii}}{d_{ii} + \sigma_{ii}^k}, \; i = 1, \ldots, n, \end{aligned}$$

[Bellavia, De Simone, di Serafino, M. 2012].

Diagonally Updated ILU (DU_ILU) [Bellavia, Morini, Porcelli 2013]

Let $\mathcal{P}_{seed} = LDU$.

1. Consider

$$J_k \simeq LDU + \underbrace{diag(J_k - J_{seed})}_{\Sigma_k = diag(\sigma_{11}^k, \ldots, \sigma_{nn}^k)}$$

2. Form the approximate factorization $\mathcal{P}_k = L_k D_k U_k$ for $LDU + \Sigma_k$

$$\begin{aligned} D_k &= D + \Sigma_k, \\ L_k &= eye(n), \quad off(L_k) = off(L) Z_k \\ U_k &= eye(n), \quad off(U_k) = Z_k off(U) \\ Z_k &= diag(z_{11}^k, \ldots, z_{nn}^k), \quad z_{ii}^k = \frac{|d_{ii}|}{|d_{ii}| + |\sigma_{ii}^k|}, \; i = 1, \ldots, n \end{aligned}$$

[Gill, Murray, Ponceleon, Saunders 1992].

Properties of DU_ILU

Scaling matrix $Z_k = diag(z_{11}^k, \ldots, z_{nn}^k)$:

$$z_{ii}^k = \frac{|d_{ii}|}{|d_{ii}| + |\sigma_{ii}^k|}, \quad i = 1, \ldots, n,$$

- Since $z_{ii}^k \in (0, 1]$, the conditioning of L_k and U_k is at least as good as the conditioning of L and U respectively [Lemeire 1975].
- If the entries of Σ_k are small then $LDU + \Sigma_k$ is close to LDU and Z_k is close to the identity matrix.

Properties of DU_ILU (c.ed)

- Quality of DU_ILU preconditioner

$$\|J_k - \mathcal{P}_k\| \leq \|J_{seed} - \mathcal{P}_{seed}\| + \|off(J_k - J_{seed})\| + c\|\Sigma_k\|$$

  The upper bound depends on

  - $\|J_{seed} - \mathcal{P}_{seed}\|$: quality of the seed preconditioner;
  - $\|off(J_k - J_{seed})\|$: information discarded in the update;
  - $\|off(J_k - J_{seed})\|$ and $\|\Sigma_k\|$ small for slowly varying sequences.

- In order to form Σ_k, $diag(J_k)$ can be evaluated by finite differences.

  If the cost for evaluating the n components of F is roughly the cost of one full F-evaluation (i.e. F is separable) then forming Σ_k amounts to one F-evaluation.

Preconditioned Newton-Krylov method with linesearch: numerical comparison under Matlab

- Linear solver: BiCGSTAB, $\mathtt{LI}_{\mathtt{max}} = 400$
- Refresh: if the backtracking strategy fails in producing an acceptable step then a J_{seed} and $\mathcal{P}_{seed}$ are initialized.
- Finite difference approximation for computing J_{seed}, J_k times a vector, $diag(J_k)$.
- Test Problems:
  - Nonlinear Convection-Diffusion ($\mathtt{NCD}$), $Re = 750, 1000, 1250$
  - Flow in a Porous Medium ($\mathtt{FPM}$),
  - CounterCurrent Reactor ($\mathtt{CCR}$)

  Varying dimension $n = 4900, 8100, 10000, 15625, 22500$.

Nonlinear Convection-Diffusion problem

Re	n	L_IT	NL_IT	Time	N_REFR
750	4900	1159	16	29.84	1
	8100	1010	15	74.45	2
	10000	1009	16	107.01	2
	15625	923	15	194.83	1
	22500	810	15	408.8	1
1000	4900	1204	16	34.54	2
	8100	799	16	71.17	2
	10000	1010	16	106.99	2
	15625	675	17	238.61	2
	22500	1281	15	423.13	1
1250	4900	877	16	31.96	2
	8100	1002	17	74.99	2
	10000	909	16	106.71	2
	15625	1068	17	256.65	2
	22500	753	17	518.49	2

Comparison between DU_ILU and Duintjer Tebbens & Tuma approach

- In terms of computational time, the DU_ILU strategy resulted to be faster than the procedure by Duintjer Tebbens and Tuma.

- In terms of F-evaluations, the two strategies were comparable (measure relevant in a nearly matrix-free setting).

Future work: *combination of an incremental factorization and an updating strategy* for sequence of symmetric indefinite matrices in constrained and unconstrained optimization.

Sequences of KKT matrices

Let $\mathcal{A}_k$ be the KKT matrix of the form

$$\mathcal{A}_k = \begin{bmatrix} Q + \Theta_k^{(1)} & A^T \\ A & -\Theta_k^{(2)} \end{bmatrix}$$

with

- $Q \in \mathbb{R}^{n \times n}$ symmetric positive semidefinite,
- $A \in \mathbb{R}^{m \times n}$, $0 < m \leq n$, full rank
- $\Theta_k^{(1)} \in \mathbb{R}^{n \times n}$ diagonal SPD,
- $\Theta_k^{(2)} \in \mathbb{R}^{m \times m}$ diagonal positive semidefinite.

This matrix arises at the kth iteration of an IP method for the convex QP problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x^T Q x + c^T x, \\ \text{s. t.} & A_1 x - s = b_1, \;\; A_2 x = b_2, \;\; x + v = u, \;\; (x, s, v) \geq 0, \end{array}$$

see e.g. [S. Wright, 1997], [D'Apuzzo, De Simone, di Serafino 2010], [Gondzio 2012]

Constraint Preconditioners (CPs)

$$\mathcal{P}_k = \begin{bmatrix} H_k & A^T \\ A & -\Theta_k^{(2)} \end{bmatrix}$$

- H_k "simple" symmetric approximation to $Q + \Theta_k^{(1)}$; here $H_k = diag(Q + \Theta_k^{(1)})$, [Benzi, Golub, Liesen 2005]

- Spectral properties of $\mathcal{P}_k^{-1}\mathcal{A}_k$. With $p = rank(\Theta_k^{(2)})$,

  an eigenvalue at 1 with multiplicity $2m - p$;

  $n - m + p$ real positive eigenvalues such that the better H_k approximates $Q + \Theta_k^{(1)}$ the more clustered around 1 they are,
  [Keller, Gould, Wathen, 2000; Dollar, 2007]

Factorization of CPs

1. Bunch-Parlett factorization

$$\mathcal{P}_k \;=\; \bar{L}_k \bar{D}_k \bar{L}_k^T,$$

$\bar{L}_k$ unit lower triang., $\bar{D}_k$ symm. block diagonal with 1×1 or 2×2 blocks.

2. Factorize the negative Schur complement S_k of H_k in $\mathcal{A}_k$

$$S_k = AH_k^{-1}A^T + \Theta_k^{(2)} = L_k D_k L_k^T \quad \text{Cholesky-like factorization}$$

and let

$$\begin{aligned}
\mathcal{P}_k &= \begin{bmatrix} I_n & 0 \\ AH_k^{-1} & I_m \end{bmatrix} \begin{bmatrix} H_k & 0 \\ 0 & -S_k \end{bmatrix} \begin{bmatrix} I_n & H_k^{-1}A^T \\ 0 & I_m \end{bmatrix} \\
&= \begin{bmatrix} I_n & 0 \\ AH_k^{-1} & L_k \end{bmatrix} \begin{bmatrix} H_k & 0 \\ 0 & -D_k \end{bmatrix} \begin{bmatrix} I_n & H_k^{-1}A^T \\ 0 & L_k^T \end{bmatrix},
\end{aligned}$$

In large-scale problems, the factorization of CPs may still account for a large part of the cost of the IP iterations.

Inexact CPs

- Approximations of CPs: based on approximate factorizations of the Schur complement or on sparse approximations of A
  [Lukšan, Vlček, 1998], [Perugia, Simoncini 2000], [Durazzi, Ruggiero 2002], [Bergamaschi, Gondzio, Venturin, Zilli, 2007].

  No exploitation of CPs for previous matrices in the sequence.

- Our focus is on inexact CPs of the form

$$(\mathcal{P}_k)_{inex} = \begin{bmatrix} I_n & 0 \\ AH_k^{-1} & I_m \end{bmatrix} \begin{bmatrix} H_k & 0 \\ 0 & -(S_k)_{inex} \end{bmatrix} \begin{bmatrix} I_n & H_k^{-1}A^T \\ 0 & I_m \end{bmatrix}$$

  where

  - $(S_k)_{inex}$ is a SPD matrix;
  - $(S_k)_{inex}$ is computationally cheaper than S_k.

Inexact CPs built by updating

1. Given

$$\mathcal{A}_{seed} = \begin{bmatrix} Q + \Theta^{(1)}_{seed} & A^T \\ A & -\Theta^{(2)}_{seed} \end{bmatrix}$$

$$S_{seed} = AH^{-1}A^T + \Theta^{(2)}_{seed} = LDL^T$$

$$\mathcal{P}_{seed} = \begin{bmatrix} I_n & 0 \\ AH^{-1} & I_m \end{bmatrix} \begin{bmatrix} H & 0 \\ 0 & -S_{seed} \end{bmatrix} \begin{bmatrix} I_n & H^{-1}A^T \\ 0 & I_m \end{bmatrix} \quad \text{seed CP}$$

2. Let

$$\mathcal{A} = \begin{bmatrix} Q + \Theta^{(1)} & A^T \\ A & -\Theta^{(2)} \end{bmatrix}, \qquad G = diag(Q + \Theta^{(1)})$$

$$S = AG^{-1}A^T + \Theta^{(2)}$$

Form an inexact CP where S is replaced by a SPD matrix obtained by updating S_{seed}.

Spectral analysis of Inexact CPs

Our updating strategy is guided by the spectral analysis for inexact CPs.

- Spectral characterization of inexact CPs has been carried out [Benzi, Simoncini, 2006], [Bergamaschi, 2011], [Sesana, Simoncini, 2013].
- The eigenvalues of the preconditioned matrix may fail to explain the behaviour of a nonsymmetric solver, e.g. when the condition number of the eigenvector matrix is far from one or when the matrix itself is higly non-normal [Greenbaum, Ptak, Strakos 1998], [Arioli, Ptak, Strakos 1998].
- Nonetheless, in many practical cases the convergence of a Krylov method applied to the preconditioned system is determined by the distribution of eigenvalues of $\mathcal{P}_k^{-1}\mathcal{A}_k$.

[Benzi, Simoncini 2006]

$\mathcal{P}_{inex}^{-1}\mathcal{A}$ has at most $2m$ eigenvalues with nonzero imaginary part, counting conjugates.

$$S_{inex} = RR^T, \quad \mathcal{P}_{inex}^{-1}\mathcal{A}w = \lambda w$$

$$\Downarrow$$

$$\begin{bmatrix} X & Y \\ Y^T & -Z \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} I_n & 0 \\ 0 & -I_m \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

where

$$\begin{aligned} X &= G^{-\frac{1}{2}}(Q + \Theta^{(1)})G^{-\frac{1}{2}} \\ Y &= (I_n - X)G^{-\frac{1}{2}}A^T R^{-T}, \\ Z &= R^{-1}\left(AG^{-\frac{1}{2}}(2I_n - X)G^{-\frac{1}{2}}A^T + \Theta^{(2)}\right)R^{-T}, \end{aligned}$$

$$
\begin{aligned}
X &= G^{-\frac{1}{2}}(Q+\Theta^{(1)})G^{-\frac{1}{2}} \\
Y &= (I_n - X)G^{-\frac{1}{2}}A^T R^{-T} \\
Z &= R^{-1}\left(AG^{-\frac{1}{2}}(2I_n - X)G^{-\frac{1}{2}}A^T + \Theta^{(2)}\right)R^{-T}
\end{aligned}
$$

If Y is full rank and Z is positive semidefinite:

- if $Im(\lambda) \neq 0$, then

$$
\begin{aligned}
&|Im(\lambda)| \leq \|Y\| \\
&\frac{1}{2}(\lambda_{min}(X) + \lambda_{min}(Z)) \leq Re(\lambda) \leq \frac{1}{2}(\lambda_{max}(X) + \lambda_{max}(Z))
\end{aligned}
$$

- if $Im(\lambda) = 0$, then either

$$\lambda_{min}(X) \leq \lambda \leq \lambda_{max}(X), \quad \text{for} \quad v = 0,$$

or

$$2\min\{\lambda_{min}(X),\, \lambda_{min}(Z)\} \leq \lambda \leq \max\{\lambda_{max}(X),\, \lambda_{max}(Z)\} \quad \text{for} \quad v \neq 0.$$

New bounds on Y and Z

The quality of S_{inex} with respect to S affects the bound on $\|Y\|$ and the spectrum of Z.

1. $\|Y\| \leq \|I_n - X\| \sqrt{\lambda_{max}(S_{inex}^{-1}\, S)}$
2. If Z is positive definite, then

$$
\begin{aligned}
\lambda_{max}(Z) &\leq \lambda_{max}(S_{inex}^{-1}\, S) \max\{2 - \lambda_{min}(X), 1\}, \\
\lambda_{min}(Z) &\geq \lambda_{min}(S_{inex}^{-1}\, S) \min\{2 - \lambda_{max}(X), 1\}.
\end{aligned}
$$

Specific choices of S_{inex} along with these bounds yield algorithmic consequences.

Building S_{upd}. Zero (2,2) block

$$S_{seed} = AH^{-1}A^T = LDL^T, \quad S = AG^{-1}A^T,$$
$$S_{upd} = AJ^{-1}A^T$$

Let $\gamma_i(J)$ be the diagonal entries of JG^{-1} sorted in nondecreasing order

$$\min_{1\le i\le n} \frac{J_{ii}}{G_{ii}} \equiv \gamma_1(J) \le \gamma_2(J) \le \cdots \le \gamma_n(J) \equiv \max_{1\le i\le n} \frac{J_{ii}}{G_{ii}}.$$

Then the eigenvalues of $S_{upd}^{-1}S$ satisfy

$$\gamma_1(J) \le \lambda(S_{upd}^{-1}S) \le \gamma_n(J).$$

$$\Downarrow$$

Form S_{upd} as a low-rank correction of S_{seed} [Baryamureeba, Steihaug, Zhang 1999]

Choosing J

1. Compute $\gamma_i(H)$ ($\gamma_1(H)$ and $\gamma_n(H)$ bounds on $\lambda(S_{seed}^{-1}S)$).
2. Let
   $\Gamma = \{\text{indices of the } q_1 \text{ largest and } q_2 \text{ smallest } \gamma_i(H)\text{'s}\}, \quad q = q_1 + q_2 \ll m$
3. Set $S_{upd} = AJ^{-1}A^T, \qquad J_{ii} = \begin{cases} G_{ii} & \text{if } i \in \Gamma \\ H_{ii} & \text{otherwise} \end{cases}$

$$\Downarrow$$

$$\begin{aligned} \gamma_n(J) &= \max\left\{1, \max_{j \notin \Gamma} \gamma_j(H)\right\} \leq \gamma_{n-q_1}(H) \\ \gamma_1(J) &= \min\left\{1, \min_{j \notin \Gamma} \gamma_j(H)\right\} \geq \gamma_{q_2+1}(H) \end{aligned}$$

Improved bounds on the eigenvalues if $\gamma_{n-q_1+1}(H)$ and $\gamma_{q_2}(H)$ are well separated from $\gamma_{n-q_1}(H)$ and $\gamma_{q_2+1}(H)$.

$\mathcal{P}_{upd}^{-1}\mathcal{A}$ has $2q$ unit eigenvalues with geometric multiplicity q

[Sesana, Simoncini 2013].

$$S_{upd} = AJ^{-1}A^T, \qquad J_{ii} = \begin{cases} G_{ii} & \text{if } i \in \Gamma \\ H_{ii} & \text{otherwise} \end{cases}$$

with $\Gamma = \{\text{indices of the } q_1 \text{ largest and } q_2 \text{ smallest } \gamma_i(H)\text{'s}\}, \quad q = q_1 + q_2 \ll m$

$$\Downarrow$$

$$S_{upd} = AJ^{-1}A^T = S_{seed} + \bar{A}\bar{K}\bar{A}^T$$

$\bar{K} \in \mathbb{R}^{q \times q}$ diagonal with entries $G_{ii}^{-1} - H_{ii}^{-1}$, $i \in \Gamma$; $\bar{A} \in \mathbb{R}^{m \times q}$ corresp. cols of A

$$\mathcal{P}_{upd} = \begin{bmatrix} I_n & 0 \\ AG^{-1} & I_m \end{bmatrix} \begin{bmatrix} G & 0 \\ 0 & -S_{upd} \end{bmatrix} \begin{bmatrix} I_n & G^{-1}A^T \\ 0 & I_m \end{bmatrix}$$

Building $\mathcal{P}_{upd}$

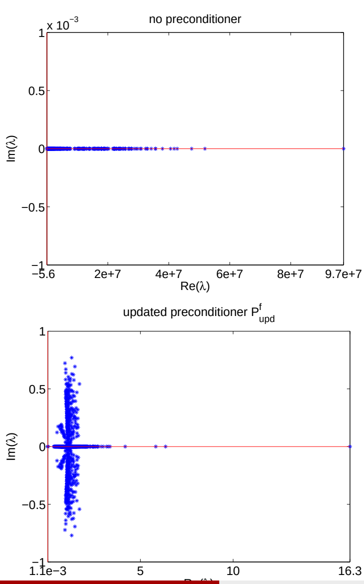
- S_{upd}^{-1} can be computed by the Sherman-Morrison formula and this yields the factorization of $\mathcal{P}_{upd}^{-1}$.
- The factorization of S_{upd} can be computed using procedures for updating/downdating a Cholesky factorization
  [Gill, Golub, Murray, Saunders, 1974], [Davis, Hager 1999, 2001, 2009].
- Assume $q = 0$, Γ empty set. Then the S_{seed} is frozen.

$$S_{upd} = S_{seed} = AH^{-1}A^T, \qquad S = AG^{-1}A^T$$

If $\left\| \Theta^{(1)} - \Theta^{(1)}_{seed} \right\|$ is small, then $\gamma_1(H)$ and $\gamma_n(H)$ are close to 1

Spectra of preconditioned matrices

CVXQP1 ($n = 1000$, $m = 5000$)
$k =$ PRQP it $= 10$, seed it $= 6$, $q = 50$



Building S_{upd}, nonzero (2,2) block

Let $\tilde{\Theta}^{(2)}$ and $\tilde{\Theta}^{(2)}_{seed}$ be the submatrix of $\Theta^{(2)}$ and $\Theta^{(2)}_{seed}$ with nonzero diagonal elements.

Let $\tilde{I}_m$ consist of all columns of I_m with indices as $\tilde{\Theta}^{(2)}$.

$$
\begin{aligned}
S_{seed} &= AH^{-1}A^T + \Theta^{(2)}_{seed} = \begin{bmatrix} A & \tilde{I}_m \end{bmatrix} \begin{bmatrix} H^{-1} & 0 \\ 0 & \tilde{\Theta}^{(2)}_{seed} \end{bmatrix} \begin{bmatrix} A^T \\ \tilde{I}_m \end{bmatrix} \\
S &= AG^{-1}A^T + \Theta^{(2)} = \begin{bmatrix} A & \tilde{I}_m \end{bmatrix} \begin{bmatrix} G^{-1} & 0 \\ 0 & \tilde{\Theta}^{(2)} \end{bmatrix} \begin{bmatrix} A^T \\ \tilde{I}_m \end{bmatrix}
\end{aligned}
$$

We consider

$$
S_{upd} = AJ^{-1}A^T + \Theta^{(2)}_{upd}
$$

where J is a low-rank update of H and $\Theta^{(2)}_{upd}$ is low-rank update of $\Theta^{(2)}_{seed}$.

$$S_{upd} = AJ^{-1}A^T + \Theta^{(2)}_{upd} = \left[\begin{array}{cc} A & \tilde{I}_m \end{array}\right] \left[\begin{array}{cc} J^{-1} & 0 \\ 0 & \tilde{\Theta}^{(2)}_{upd} \end{array}\right] \left[\begin{array}{c} A^T \\ \tilde{I}_m \end{array}\right]$$

If $\gamma_i(J, \tilde{\Theta}^{(2)}_{upd})$ are the diagonal entries of

$$\left[\begin{array}{cc} J & 0 \\ 0 & (\tilde{\Theta}^{(2)}_{upd})^{-1} \end{array}\right] \left[\begin{array}{cc} G^{-1} & 0 \\ 0 & \tilde{\Theta}^{(2)} \end{array}\right]$$

sorted in nondecreasing order,

$$\gamma_1(J, \tilde{\Theta}^{(2)}_{upd}) \le \gamma_2(J, \tilde{\Theta}^{(2)}_{upd}) \le \cdots \le \gamma_n(J, \tilde{\Theta}^{(2)}_{upd})$$

Then the eigenvalues of $S_{upd}^{-1}S$ satisfy

$$\gamma_1(J, \tilde{\Theta}^{(2)}_{upd}) \le \lambda(S_{upd}^{-1}S) \le \gamma_n(J, \tilde{\Theta}^{(2)}_{upd}).$$

$$\Downarrow$$

Form S_{upd} as a low-rank correction of S

J and $\Theta^{(2)}_{upd}$ low-rank corrections of H and $\Theta^{(2)}_{seed}$ built as before

Preliminary numerical experiments: sequences from PRQP

PRQP

- Fortran 90 primal-dual Potential Reduction solver for convex Quadratic Programming (feasible and infeasible versions)
  [Cafieri, D'Apuzzo, De Simone, di Serafino, Toraldo, 2007-2010]

Preliminary numerical experiments: sequences from PRQP

PRQP

- Fortran 90 primal-dual Potential Reduction solver for convex Quadratic Programming (feasible and infeasible versions)
  [Cafieri, D'Apuzzo, De Simone, di Serafino, Toraldo, 2007-2010]

Test KKT sequences

- Sequences of KKT systems (with right-hand sides and CG tolerances) obtained by running PRQP on convex QP problems with linear equality constraints only ($\Theta^{(2)} = 0$)
- QP problems from CUTEr (different size and structure of the KKT matrix and its Schur complement)

Preliminary numerical experiments: implementation details

- Solution of KKT systems by SQMR (Matlab implementation)
- Sparse LDL^T factorizations and updates/downdates computed by using CHOLMOD through its Matlab interface [Davis, Hager et al., 2008-2009]
- $q = q_1 + q_2 = 0, 50, 100$, only $\gamma_i(H) = H_{ii}/G_{ii} > 10$ or $\gamma_i(H) = H_{ii}/G_{ii} < 0.1$
- "Refresh" of $\mathcal{P}_{seed}$
  - after a fixed max number of updates has been performed;
  - when the time (prec + solve) for solving a system exceeds 90% of the time for the last solution with the CP.
- Computational environment: Intel Core i7 processor @ 2.67 GHz, 12 GB of RAM and 8 MB of cache memory; Linux O.S. 3.2.0-35-generic; gcc 4.3.4 compiler; Matlab R2011b (7.13); CHOLMOD 2.1.2

Comparison of $\mathcal{P}_k$ with $\mathcal{P}_{upd}$

Problem	n, m	constr. prec		P_{upd} $(q=0)$		P_{upd} $(q=50)$		P_{upd} $(q=100)$	
		nit	*time*	*nit*	*time*	*nit*	*time*	*nit*	*time*
CVXQP1	20000 10000	232	5.76e+1	740	3.79e+1	530	3.72e+1	495	4.05e+1
CVXQP3	20000 15000	497	8.32e+2	2006	4.82e+2	1166	3.67e+2	1118	3.76e+2
CVXQP2	20000 5000	273	1.29e+0	352	1.65e+0	330	1.57e+0	308	1.46e+0
STCQP2	16385 8190	259	1.44e+0	267	1.47e+0	267	1.49e+0	267	1.49e+0

nit = # SQMR iterations, max # updates: 4

CVXQP1 ($n = 20000, m = 10000$): details

k	exact constr. prec.			P_{upd} (q=50)		
	nit	T_{prec}	T_{solve}	nit	T_{prec}	T_{solve}
1	29	2.75e+0	9.87e-1	29	2.75e+0	9.87e-1
2	16	2.74e+0	5.50e-1	19	3.76e-1	6.60e-1
3	10	2.73e+0	3.60e-1	14	6.87e-1	4.91e-1
4	6	2.75e+0	2.22e-1	18	6.54e-1	6.18e-1
5	4	2.75e+0	1.56e-1	24	6.68e-1	8.36e-1
6	6	2.74e+0	2.21e-1	6	2.74e+0	2.21e-1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
11	12	3.02e+0	4.46e-1	12	3.02e+0	4.46e-1
12	13	3.01e+0	4.79e-1	19	1.04e-1	7.08e-1
13	14	3.01e+0	5.19e-1	27	6.48e-1	9.40e-1
14	16	3.01e+0	5.84e-1	74	6.51e-1	2.59e+0
15	18	3.01e+0	6.51e-1	18	3.01e+0	6.51e-1
16	18	3.03e+0	6.57e-1	25	2.15e-1	9.19e-1
17	34	3.05e+0	1.24e+0	106	5.76e-1	3.78e+0
	232	4.92e+1	8.41e+0	530	1.84e+1	1.88e+1
		5.76e+1			3.72e+1	

k = PRQP iteration, nit = # SQMR iterations

blue: prec. refresh, red: column sum

STCQP2 ($n = 20000$, $m = 10000$): details

k	exact constr. prec.			P_{upd} (q=50)		
	nit	T_{prec}	T_{solve}	nit	T_{prec}	T_{solve}
1	13	1.08e-2	8.78e-2	13	1.08e-2	8.78e-2
2	11	1.01e-2	5.88e-2	13	3.83e-3	7.05e-2
3	11	1.01e-2	5.89e-2	12	3.78e-3	6.44e-2
4	11	1.02e-2	6.00e-2	10	3.86e-3	5.40e-2
5	11	1.01e-2	6.00e-2	13	3.84e-3	6.76e-2
6	15	1.00e-2	7.68e-2	15	1.00e-2	7.68e-2
7	19	1.01e-2	9.73e-2	22	4.07e-3	1.20e-1
8	22	1.01e-2	9.48e-2	25	4.10e-3	1.38e-1
9	24	1.01e-2	1.19e-1	33	4.10e-3	1.81e-1
10	24	1.00e-2	1.17e-1	34	3.94e-3	1.82e-1
⋮	⋮	⋮	⋮	⋮	⋮	⋮

k = PRQP iteration, nit = # SQMR iterations

blue: prec. refresh

Work in progress and open issues

- Experiments for the case $\Theta^{(2)} \neq 0$ are in progress.
- Adaptive strategies for choosing between the CP and the updated Inexact CP.
- Further analysis of the convergence of optimal Krylov solver with $\mathcal{P}_{upd}$.
- Proposal of alternative strategies for defining an updated approximate Schur complement S_{upd}.
- Study of updating techniques for the Bunch-Parlett factorization.

References

S. Bellavia, D. Bertaccini, B. Morini, *Nonsymmetric preconditioner updates in Newton-Krylov methods for nonlinear systems*, SIAM J. Sci. Comput., 2011.

S. Bellavia, B. Morini, M. Porcelli, *New updates of incomplete LU factorizations and applications to large nonlinear systems*, Optimization Methods and Software, 2013.

M. Benzi, M. Tůma, *A sparse approximate inverse preconditioner for nonsymmetric linear systems*, SIAM J. Sci. Comput., 1998.

M. Benzi, M. Tůma, *A Comparative Study of Sparse Approximate Inverse Preconditioners*, Appl. Numer. Math., 1999.

References

V. Baryamureeba, T. Steihaug, Y. Zhang, *Properties of a Class of Preconditioners for Weighted Least Squares Problems*, 1999
Technical Report No. 170, Department of Informatics, University of Bergen
Technical Report No. TR99-16, Department of Computational and Applied Mathematics, Rice University, Houston.

M. Benzi, V. Simoncini, *On the eigenvalues of a class of saddle point matrices*, Numer. Math., 2006.

D. Sesana, V. Simoncini, *Spectral analysis of inexact constraint preconditioning for symmetric saddle point matrices*, Linear Algebra and Appl., 2013.

W.Wang, D. P. O'Leary, *Adaptive use of iterative methods in predictor-corrector interior point methods for linear programming*, Numerical Algorithms, 2000.