Notes on explicit algebraic preconditioning (Michele, 10/17/2013)

Let $\mathcal{A}$ be a large, sparse, nonsingular $n \times n$ matrix, partitioned as

$$\mathcal{A} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Here A is $p \times p$ and D is $q \times q$, with $p+q=n$. The blocks B and C are generally rectangular. We assume that A is nonsingular.

Define the "ideal" block triangular preconditioner $\mathcal{P}$ as

$$\mathcal{P} = \begin{bmatrix} A & 0 \\ C & S \end{bmatrix},$$

where $S = D - CA^{-1}B^T$ is the Schur complement. Note that S is necessarily nonsingular, since

$$\det(\mathcal{A}) = \det(A)\det(S).$$

We have

$$\mathcal{P}^{-1}\mathcal{A} = \begin{bmatrix} I_p & A^{-1}B \\ 0 & I_q \end{bmatrix},$$

hence

$$H = I_n - \mathcal{P}^{-1}\mathcal{A} = \begin{bmatrix} 0 & -A^{-1}B \\ 0 & 0 \end{bmatrix},$$

which has only one nonzero block, of order $p \times q$; if q is small relative to p, or if A^{-1} is sparse, then the only nonzero block in H is sparse. Note that $\rho(H) = 0$ and therefore $H^2 = 0$: the Neumann series reduces to $I + H$.

As a very simple example, take the $2D$ Laplacian or ADR operator on the unit square, discretized with finite differences (5-point stencil). Then a red-black ordering of the grid points produces the matrix

$$\mathcal{A} = \begin{bmatrix} E & B \\ C & E \end{bmatrix}.$$

with D diagonal and B, C square of the same size as E. Now, E^{-1} is diagonal and

$$H = \begin{bmatrix} 0 & -E^{-1}B \\ 0 & 0 \end{bmatrix},$$

an extremely sparse matrix. Furthermore, $S = E - CE^{-1}B$ is also quite sparse.

As another example, consider a non-overlapping domain decomposition (or graph partitioning) approach using K subdomains. The coefficient matrix can be reordered as

$$\mathcal{A} = \begin{bmatrix} A_1 & & & & B_1 \\ & A_2 & & & B_2 \\ & & \ddots & & \vdots \\ & & & A_K & B_K \\ C_1 & C_2 & \cdots & C_K & D \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Using

$$\mathcal{P} = \begin{bmatrix} A_1 & & & & B_1 \\ & A_2 & & & B_2 \\ & & \ddots & & \vdots \\ & & & A_K & B_K \\ 0 & 0 & \cdots & 0 & S \end{bmatrix}$$

(with $S = D - CA^{-1}B = D - \sum_{i=1}^{K} C_i A_i^{-1} B_i$) as the "ideal" preconditioner we obtain

$$H = I - \mathcal{P}^{-1}\mathcal{A} = \begin{bmatrix} 0 & & & & -A^{-1}B_1 \\ & 0 & & & -A_2^{-1}B_2 \\ & & \ddots & & \vdots \\ & & & 0 & -A_K^{-1}B_K \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

and again $H^2 = 0$. Note that H is fairly sparse, especially if the separator set (= the dimension of the matrix D) is not too large.

In practice, these preconditioners are often too expensive due to the need to solve linear systems involving the Schur complement S at each iteration. Instead, "exact" solves with S are replaced with "approximate solves" with an approximate Schur complement $\hat{S} \approx S$. (Note: Solves involving the block A may also be approximated but here we assume they are performed "exactly"). Using the preconditioner

$$\hat{\mathcal{P}} = \begin{bmatrix} A & 0 \\ C & \hat{S} \end{bmatrix}$$

produces the preconditioned matrix

$$\hat{\mathcal{P}}^{-1}\mathcal{A} = \begin{bmatrix} I & A^{-1}B \\ 0 & \hat{S}^{-1}S \end{bmatrix},$$

and therefore the iteration matrix

$$\hat{H} = I - \hat{\mathcal{P}}^{-1}\mathcal{A} = \begin{bmatrix} 0 & -A^{-1}B \\ 0 & I - \hat{S}^{-1}S \end{bmatrix}.$$

Note that this matrix is not as sparse as in the "ideal" case but the amount of fill could be acceptable, depending on the choice of $\hat{S}$. Moreover,

$$\rho(\hat{H}) = \rho(I - \hat{S}^{-1}S)$$

and therefore the approximation $\hat{S} \approx S$ must be such that $\rho(I - \hat{S}^{-1}S) < 1$.

For example, one can prove that if $\mathcal{A}$ is SPD (symmetric positive definite), so are S and D and moreover using just D to approximate $S = D - B^T A^{-1} B$ results in a spectral radius less than 1 and an iteration matrix

$$\hat{H} = I - \hat{\mathcal{P}}^{-1}\mathcal{A} = \begin{bmatrix} 0 & -A^{-1}B \\ 0 & D^{-1}CA^{-1}B \end{bmatrix}.$$

Now the spectral radius is $\neq 0$ and the Neumann series $I + \hat{H} + \hat{H}^2 + \ldots$ is infinite.

The (2,2) block in $\hat{H}$ could be fairly dense, but other choices of D leading to sparser $\hat{H}$ may exist.