A framework for deflated BiCG and related solvers

Martin H. Gutknecht
Seminar for Applied Mathematics, ETH Zurich
http://www.sam.math.ethz.ch/~mhg

SIAM Conference on Applied Linear Algebra
Valencia, 20 June 2012

Joint work with André Gaul, Jörg Liesen, Reinhard Nabben
Outline

Augmentation and Deflation: Basics

History

Galerkin: CG, CR, GCR, ...

Petrov-Galerkin: BiCG, BiCR

Conclusions
Iterative methods based on (Petrov-)Galerkin conditions

To solve: \( Ax = b \) with \( A \in \mathbb{C}^{N\times N} \) nonsingular.

Construct sequence \( x_n \) such that \( r_n \equiv b - Ax_n \to 0 \).
Choose \( x_n \) from search space \( x_0 + S_n \) such that some Galerkin or Petrov-Galerkin condition is satisfied:

\[
x_n \in x_0 + S_n, \quad r_n = A(x_\star - x_n) \perp B^H \tilde{S}_n
\]

with some (formal) inner product matrix \( B^H \), i.e.,

\[
r_n \in r_0 + AS_n, \quad r_n \perp B^H \tilde{S}_n.
\]

\[\Leftrightarrow \] \( r_0 \) is approximated from \( AS_n \) such that “error” \( r_n \perp \tilde{S}_n \).
Iterative methods based on (Petrov-)Galerkin conditions

To solve: \( \mathbf{Ax} = \mathbf{b} \) with \( \mathbf{A} \in \mathbb{C}^{N \times N} \) nonsingular.

Construct sequence \( \mathbf{x}_n \) such that \( \mathbf{r}_n :\equiv \mathbf{b} - \mathbf{Ax}_n \to \mathbf{0} \).

Choose \( \mathbf{x}_n \) from search space \( \mathbf{x}_0 + \mathbf{S}_n \) such that some Galerkin or Petrov-Galerkin condition is satisfied:

\[
\mathbf{x}_n \in \mathbf{x}_0 + \mathbf{S}_n, \quad \mathbf{r}_n = \mathbf{A}(\mathbf{x}^* - \mathbf{x}_n) \perp \mathbf{B}^\mathbf{H}\tilde{\mathbf{S}}_n
\]

with some (formal) inner product matrix \( \mathbf{B}^\mathbf{H} \), i.e.,

\[
\mathbf{r}_n \in \mathbf{r}_0 + \mathbf{A}\mathbf{S}_n, \quad \mathbf{r}_n \perp \mathbf{B}^\mathbf{H}\tilde{\mathbf{S}}_n.
\]

\( \leadsto \mathbf{r}_0 \) is approximated from \( \mathbf{A}\mathbf{S}_n \) such that “error” \( \mathbf{r}_n \perp \tilde{\mathbf{S}}_n \).

Two cases: Galerkin: \( \tilde{\mathbf{S}}_n = \mathbf{S}_n \), Petrov-Galerkin: \( \tilde{\mathbf{S}}_n \neq \mathbf{S}_n \)
Iterative methods based on (Petrov-)Galerkin conditions

To solve: \( \mathbf{Ax} = \mathbf{b} \) with \( \mathbf{A} \in \mathbb{C}^{N \times N} \) nonsingular.

Construct sequence \( \mathbf{x}_n \) such that \( \mathbf{r}_n :\equiv \mathbf{b} - \mathbf{Ax}_n \to \mathbf{0} \).

Choose \( \mathbf{x}_n \) from search space \( \mathbf{x}_0 + \mathcal{S}_n \) such that some Galerkin or Petrov-Galerkin condition is satisfied:

\[
\mathbf{x}_n \in \mathbf{x}_0 + \mathcal{S}_n, \quad \mathbf{r}_n = \mathbf{A}(\mathbf{x}_* - \mathbf{x}_n) \perp \mathbf{B}^\mathbf{H} \tilde{\mathcal{S}}_n
\]

with some (formal) inner product matrix \( \mathbf{B}^\mathbf{H} \), i.e.,

\[
\mathbf{r}_n \in \mathbf{r}_0 + \mathbf{A}\mathcal{S}_n, \quad \mathbf{r}_n \perp \mathbf{B}^\mathbf{H} \tilde{\mathcal{S}}_n.
\]

\( \mathbf{r}_0 \) is approximated from \( \mathbf{A}\mathcal{S}_n \) such that “error” \( \mathbf{r}_n \perp \tilde{\mathcal{S}}_n \).

Two cases: Galerkin: \( \mathcal{S}_n = \mathcal{S}_n \), Petrov-Galerkin: \( \mathcal{S}_n \neq \mathcal{S}_n \)

\( \mathbf{B} :\equiv \mathbf{I} \) for CG/BiCG, \( \mathbf{B} :\equiv \mathbf{A}^\mathbf{H} \) for CR/GCR, \( \mathbf{B} :\equiv \mathbf{A} \) for BiCR.
Augmentation and deflation

Ass.: know basis $\mathbf{U}$ of approximately $\mathbf{A}$-invariant subspace $\mathbf{U}$, i.e., $\mathbf{U} = \mathcal{R}(\mathbf{U})$, where $\mathbf{U} \in \mathbb{C}^{N \times k}$ full rank.

Search space $\mathcal{S}_n$ and test space $\mathcal{\tilde{S}}_n$ are split up:

$$
\mathbf{x}_n \in \mathbf{x}_0 + \mathcal{S}_n, \quad \mathbf{r}_n \perp \mathbf{B}^H \mathcal{\tilde{S}}_n,
$$

$$
\mathcal{S}_n \equiv \mathcal{\hat{K}}_n \oplus \mathbf{U}, \quad \mathcal{\tilde{S}}_n \equiv \mathcal{\hat{L}}_n \oplus \mathcal{\tilde{U}},
$$

$$
\mathcal{\hat{K}}_n \equiv \mathcal{K}_n(\mathcal{\hat{A}}, \mathcal{\hat{r}}_0) := \text{span}\{\mathcal{\hat{r}}_0, \ldots, \mathcal{\hat{A}}^{n-1} \mathcal{\hat{r}}_0\}.
$$

Still to be specified: $\mathcal{\hat{A}}, \mathcal{\hat{r}}_0, \mathcal{\hat{L}}_n, \mathcal{\tilde{U}}$. 
Augmentation and deflation

Ass.: know basis $\mathbf{U}$ of approximately $\mathbf{A}$-invariant subspace $\mathcal{U}$, i.e., $\mathcal{U} = \mathcal{R}(\mathbf{U})$, where $\mathbf{U} \in \mathbb{C}^{N \times k}$ full rank.

Search space $S_n$ and test space $\tilde{S}_n$ are split up:

$$
x_n \in \mathbf{x}_0 + S_n, \quad r_n \perp B^H \tilde{S}_n,
$$

$$
S_n \equiv \tilde{\mathcal{K}}_n \oplus \mathcal{U}, \quad \tilde{S}_n \equiv \tilde{\mathcal{L}}_n \oplus \tilde{\mathcal{U}},
$$

$$
\tilde{\mathcal{K}}_n \equiv \mathcal{K}_n(\tilde{\mathbf{A}}, \tilde{r}_0) \equiv \text{span} \{ \tilde{r}_0, \ldots, \tilde{\mathbf{A}}^{n-1}\tilde{r}_0 \}.
$$

Still to be specified: $\tilde{\mathbf{A}}, \tilde{r}_0, \tilde{\mathcal{L}}_n, \tilde{\mathcal{U}}$.

Galerkin case: $\tilde{\mathcal{L}}_n \equiv \tilde{\mathcal{K}}_n$, $\tilde{\mathcal{U}} \equiv \mathcal{U}$. 
Augmentation and deflation

Ass.: know basis $\mathbf{U}$ of approximately $\mathbf{A}$-invariant subspace $\mathbf{U}$, i.e., $\mathbf{U} = \mathcal{R}(\mathbf{U})$, where $\mathbf{U} \in \mathbb{C}^{N \times k}$ full rank.

Search space $S_n$ and test space $\tilde{S}_n$ are split up:

$$\mathbf{x}_n \in \mathbf{x}_0 + S_n, \quad \mathbf{r}_n \perp \mathbf{B}^H \tilde{S}_n,$$

$$S_n :\equiv \tilde{\mathcal{K}}_n \oplus \mathbf{U}, \quad \tilde{S}_n :\equiv \tilde{\mathcal{L}}_n \oplus \tilde{\mathcal{U}},$$

$$\tilde{\mathcal{K}}_n :\equiv \mathcal{K}_n(\tilde{\mathbf{A}}, \tilde{\mathbf{r}}_0) :\equiv \text{span} \{\tilde{\mathbf{r}}_0, \ldots, \tilde{\mathbf{A}}^{n-1}\tilde{\mathbf{r}}_0\}.$$ 

Still to be specified: $\tilde{\mathbf{A}}, \tilde{\mathbf{r}}_0, \tilde{\mathcal{L}}_n, \tilde{\mathcal{U}}.$

Galerkin case: $\tilde{\mathcal{L}}_n :\equiv \tilde{\mathcal{K}}_n, \quad \tilde{\mathcal{U}} :\equiv \mathbf{U}.$

Petrov-Galerkin case: e.g., $\tilde{\mathcal{L}}_n :\equiv \mathcal{K}_n(\tilde{\mathbf{A}}^H, \tilde{\mathbf{r}}_0)$, but other options for $\tilde{\mathcal{K}}_n$ and $\tilde{\mathcal{L}}_n$ exist.
Rationale of augmentation and deflation

Ideally: columns of $U \in \mathbb{C}^{N \times k}$ span $A$-invariant subspace $U$ belonging to eigenvalues close to 0.

Let $Z :\equiv AU$, $\mathcal{Z} :\equiv AU = U$.

Note: images of the restriction $A^{-1}|_Z$ are trivial to compute:
if $z = Zc \in \mathcal{Z}$, then $A^{-1}z = Uc$. 
Rationale of augmentation and deflation

Ideally: columns of $U \in \mathbb{C}^{N \times k}$ span $A$-invariant subspace $U$ belonging to eigenvalues close to 0.

Let $Z \equiv AU$, $\mathcal{Z} \equiv AU = U$.

Note: images of the restriction $A^{-1}|_{\mathcal{Z}}$ are trivial to compute: if $z = Zc \in \mathcal{Z}$, then $A^{-1}z = Uc$.

Choose projector $P$ such that $\mathcal{N}(P) = \mathcal{Z}$.

Split up space: $\mathbb{C}^N = \mathcal{R}(P) \oplus \mathcal{Z}$.

Choose $\hat{A} \equiv PA$ and $\hat{r}_0 \equiv Pr_0$ so that $\mathcal{K}_n \subseteq \mathcal{R}(P)$. 
Rationale of augmentation and deflation

Ideally: columns of $U \in \mathbb{C}^{N \times k}$ span $A$-invariant subspace $U$ belonging to eigenvalues close to 0.

Let $Z : = AU$, $\mathcal{Z} : = AU = U$.

Note: images of the restriction $A^{-1}|_Z$ are trivial to compute: if $z = Zc \in \mathcal{Z}$, then $A^{-1}z = Uc$.

Choose projector $P$ such that $\mathcal{N}(P) = \mathcal{Z}$.

Split up space: $\mathbb{C}^N = \mathcal{R}(P) \oplus \mathcal{Z}$.

Choose $\hat{A} : = PA$ and $\hat{r}_0 : = Pr_0$ so that $\mathcal{K}_n \subseteq \mathcal{R}(P)$.

Split up $r_0$ accordingly: $r_0 = \hat{r}_0 + (r_0 - \hat{r}_0)$.

$A^{-1}(r_0 - \hat{r}_0)$ is trivial;
$\hat{A}^{-1}\hat{r}_0$ is found with Krylov space solver acting on $\mathcal{R}(P)$.
Since \( \mathcal{N}(P) = \mathcal{Z} \), we have \( \mathcal{N}(\hat{A}) = \mathcal{N}(PA) = \mathcal{U} \).

So the (absolutely) small eigenvalues of \( A \) represented by \( U \) that caused trouble are replaced by a \( k \)-fold EVal \( o \) in \( \hat{A} \) (deflation).

Projector \( P \) is not fully determined by its nullspace since it may be oblique. Hopefully,

\[
A|_{\mathcal{R}(P)} = \hat{A}|_{\mathcal{R}(P)}.
\]  

(1)
Since $\mathcal{N}(P) = \mathcal{Z}$, we have $\mathcal{N}(\hat{A}) = \mathcal{N}(PA) = \mathcal{U}$.

So the (absolutely) small eigenvalues of $A$ represented by $U$ that caused trouble are replaced by a $k$-fold EVal $\mathbf{0}$ in $\hat{A}$ (deflation).

Projector $P$ is not fully determined by its nullspace since it may be oblique. Hopefully,

$$A|_{\mathcal{R}(P)} = \hat{A}|_{\mathcal{R}(P)}.$$  \hspace{1cm} (1)

If $A$ is Hermitian and $U$ is $A$-invariant, and if $P$ is chosen such that $\mathcal{R}(P) = \mathcal{U}^\perp$ or $\mathcal{R}(P) = \mathcal{Z}^\perp$, Eq. (1) holds.

If $\mathcal{R}(P) = \mathcal{Z}^\perp$, $P$ is an orthogonal projector.
How to find an approximately invariant subspace?

- It may be known from a theoretical analysis of the problem.
- It may result from the solution of previous systems with the same \( A \). (\( \sim \) linear system with multiple right-hand sides)
- It may result from the solution of previous systems with nearby \( A \).
- It may result from previous cycles of the solution process (if the method is restarted).

There are lots of examples in the literature.
Things to distinguish

**Augmented bases:** $x_n \in x_0 + K_n(\hat{A}, \hat{r}_0) + U$, where

$\hat{A} = A$ or $\text{spec}(\hat{A}) \subset \text{spec}(A) \cup \{0\}$

**(Spectral) deflation:** $A \rightsquigarrow \hat{A} \equiv PA \text{ s.t. small EVals} \rightsquigarrow 0$

**EVal translation:** $A \rightsquigarrow \hat{A} \equiv AP \text{ s.t. small EVals} \rightsquigarrow |\lambda_{max}|$

**Krylov space recycling:** choice of $U$ based on prev. cycles

**Flexible KSS:** adaptation of $P$ at each restart
Things to distinguish

**Augmented bases:** $x_n \in x_0 + \mathcal{K}_n(\hat{A}, \hat{r}_0) + U$, where

$\hat{A} = A$ or $\text{spec}(\hat{A}) \subset \text{spec}(A) \cup \{0\}$

**(Spectral) deflation:** $A \rightsquigarrow \hat{A} :\equiv PA$ s.t. small EVals $\rightsquigarrow 0$

**EVal translation:** $A \rightsquigarrow \hat{A} :\equiv AP$ s.t. small EVals $\rightsquigarrow |\lambda_{\max}|$

**Krylov space recycling:** choice of $U$ based on prev. cycles

**Flexible KSS:** adaptation of $P$ at each restart

Two basic approaches:

- Augmentation of basis with or without spectral deflation.
- EVal translation by suitable preconditioning (**no deflation!**).
History

Early contributions (many more papers appeared since):

Nicolaides '85/'87_{SINUM}: deflated 3-term CG (w/augm. basis)
Dostál '87/'88_{IntJCompMath}: deflated 2-term CG (w/augm. basis)
Morgan '93/'95_{SIMAX}: GMRES with augmented basis
de Sturler '93/'96_{JCAM}: inner-outer GMRES/GCR (and, briefly, inner-outer BiCGStab/GCR) with augmented basis
Chapman / Saad '95/'97_{NLAA} GMRES with augmented basis
Saad '95/'97_{SIMAX} Analysis of KSS with augmented basis
de Sturler '97/'99_{SINUM} inner-outer GMRES/GCR w/truncation
Vuik / Segal / Meijerink '98/'99_{JCP} 2-term CG w/augm. basis
Bristeau / Erhel '98/'98_{NumAlg} CG with augmented basis
Erhel / Guyomarc'h '97/'00_{SIMAX} defl. 2-term CG w/augm. basis
Saad / Yeung / Erhel / Guyomarc'h '98/'00_{SISC} the same
Galerkin: CG, CR, GCR, ... : some details

Given: \( A, B \in \mathbb{C}^{N \times k} \) and \( U \in \mathbb{C}^{N \times k} \)

Most relevant cases: \( B = I \) for CG, \( B = A^H \) for CR, GCR

\[
\begin{align*}
E &\equiv U^HBAU \in \mathbb{C}^{k \times k}, & \text{assumed nonsingular,} \\
M &\equiv UE^{-1}U^H, \\
P &\equiv I - AMB, & \text{projector onto } (B^HU)^\perp \text{ along } Z, \\
Q &\equiv I - MBA, & \text{projector onto } (A^HB^HU)^\perp \text{ along } U, \\
\hat{A} &\equiv PA = AQ = PAQ \\
\hat{r}_n &\equiv P(b - A\hat{x}_n) = Pb - \hat{A}\hat{x}_n \in (B^HU)^\perp, \\
\hat{K}_n &\equiv \mathcal{K}(\hat{A},\hat{r}_0) \subseteq (B^HU)^\perp.
\end{align*}
\]
An equivalence theorem

**Theorem**

For \( n \geq 1 \) the two pairs of conditions,

\[
\begin{align*}
\mathbf{x}_n & \in \mathbf{x}_0 + \hat{\mathcal{K}}_n + \mathcal{U}, & \mathbf{r}_n & \perp B\hat{\mathcal{K}}_n + B\tilde{\mathcal{U}}, \\
\end{align*}
\]

(2)

and

\[
\begin{align*}
\hat{\mathbf{x}}_n & \in \mathbf{x}_0 + \hat{\mathcal{K}}_n, & \hat{\mathbf{r}}_n & \perp B\hat{\mathcal{K}}_n. \\
\end{align*}
\]

(3)

are equivalent in the sense that

\[
\begin{align*}
\mathbf{x}_n = Q\hat{\mathbf{x}}_n + MB^Hb & \quad \text{and} \quad \mathbf{r}_n = \hat{\mathbf{r}}_n. \\
\end{align*}
\]

(4)

**Def.** The direct deflation approach is given by (2), the indirect deflation approach is given by (3)–(4).
Computing $\hat{x}_n$ satisfying (3) means solving the singular linear system

$$\hat{A} \hat{x} = \hat{P}b$$

with a Krylov space solver characterized by (3).
Computing $\hat{x}_n$ satisfying (3) means solving the singular linear system

$$\hat{A} \hat{x} = Pb$$

with a Krylov space solver characterized by (3).

**What are the properties of $\hat{A}$?**

$$\mathcal{N}(\hat{A}) = \mathcal{U}, \quad \mathcal{R}(\hat{A}) = (B^H U)^\perp.$$
Computing $\hat{x}_n$ satisfying (3) means solving the singular linear system

$$\hat{A} \hat{x} = Pb$$

with a Krylov space solver characterized by (3).

**What are the properties of $\hat{A}$ ?**

$$\mathcal{N}(\hat{A}) = \mathcal{U}, \quad \mathcal{R}(\hat{A}) = (B^H\mathcal{U})^\perp.$$ 

If $\mathcal{U}$ is $A$-invariant, the corresp. EVals become 0.

**What can we say about the others?**
Computing $\hat{x}_n$ satisfying (3) means solving the singular linear system

$$\hat{A} \hat{x} = Pb$$

with a Krylov space solver characterized by (3).

**What are the properties of $\hat{A}$ ?**

$$\mathcal{N}(\hat{A}) = \mathcal{U}, \quad \mathcal{R}(\hat{A}) = (B^H\mathcal{U})^\perp.$$  

If $\mathcal{U}$ is $A$-invariant, the corresp. EVals become 0.

**What can we say about the others?**

Consider partitioned Jordan decomposition of $A$

$$A = SJS^{-1} = \begin{bmatrix} S_1 & S_2 \end{bmatrix} \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} \begin{bmatrix} \hat{S}_1^H \\ \hat{S}_2^H \end{bmatrix} ,$$

where $S_1, \hat{S}_1 \in \mathbb{C}^{N \times k}$, $S_2, \hat{S}_2 \in \mathbb{C}^{N \times (N-k)}$ and either $\mathcal{R}(S_1) = \mathcal{U}$ or $\mathcal{R}(\hat{S}_1) = B^H\mathcal{U}$.  

M.H. Gutknecht  
SIAM-ALA12  
p. 12
Theorem

(1) If $\mathcal{U} = \mathcal{R}(S_1)$, if $\mathcal{U} \in \mathbb{C}^{N \times k}$ is any matrix satisfying $\mathcal{R}(\mathcal{U}) = \mathcal{U}$, and if $\mathcal{U}^H \mathcal{B} \mathcal{A} \mathcal{U}$ is nonsingular, then

$$\hat{A} = PA = \begin{bmatrix} \mathcal{U} & PS_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & J_2 \end{bmatrix} \begin{bmatrix} \mathcal{U} & PS_2 \end{bmatrix}^{-1}$$

with

$$\begin{bmatrix} \mathcal{U} & PS_2 \end{bmatrix}^{-1} = \begin{bmatrix} B^H \mathcal{U}(\mathcal{U}^H B^H \mathcal{U})^{-1} \hat{S}_2 \end{bmatrix}^H.$$

(2) If $B^H \mathcal{U} = \mathcal{R}(\hat{S}_1)$, if $\mathcal{U} \in \mathbb{C}^{N \times k}$ is any matrix satisfying $\mathcal{R}(\mathcal{U}) = \mathcal{U}$, and if $\mathcal{U}^H \mathcal{B} \mathcal{A} \mathcal{U}$ is nonsingular, then

$$\hat{A} = PA = \begin{bmatrix} \mathcal{U} & S_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & J_2 \end{bmatrix} \begin{bmatrix} \mathcal{U} & S_2 \end{bmatrix}^{-1}$$

with

$$\begin{bmatrix} \mathcal{U} & S_2 \end{bmatrix}^{-1} = \begin{bmatrix} B^H \mathcal{U}(\mathcal{U}^H B^H \mathcal{U})^{-1} Q^H \hat{S}_2 \end{bmatrix}^H.$$
Petrov-Galerkin: BICG, BICR, ...

Generalized BICG (GENBICG) [G. ’90(CopperMtn), ’97 ActaNum] with formal inner product matrix $B$ requires $A$ and $B$ to commute (to maintain short recurrences).
Petrov-Galerkin: BICG, BICR, ...

Generalized BICG (GENBICG) \([G. '90(CopperMtn), '97ActaNum]\) with formal inner product matrix \(\mathbf{B}\) requires \(\mathbf{A}\) and \(\mathbf{B}\) to commute (to maintain short recurrences).

For deflated solvers based on Petrov-Galerkin condition we need projectors and operators for creating split dual spaces:

\[ S_n := \tilde{\mathcal{K}}_n \oplus \mathcal{U}, \quad \tilde{S}_n := \tilde{\mathcal{L}}_n \oplus \tilde{\mathcal{U}}. \]

May consider solving two dual systems at once [Ahuja '09Diss]:

\[ \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{A}^H\tilde{\mathbf{x}} = \tilde{\mathbf{b}} \]

such that

\[ \mathbf{x}_n \in \mathbf{x}_0 + S_n, \quad \tilde{\mathbf{x}}_n \in \tilde{\mathbf{x}}_0 + \tilde{S}_n, \]

\[ \mathbf{r}_n \perp \mathbf{B}^H \tilde{S}_n, \quad \tilde{\mathbf{r}}_n \perp \mathbf{B}S_n. \]
Petrov-Galerkin: Projectors and other operators

<table>
<thead>
<tr>
<th></th>
<th>definition</th>
<th>null space</th>
<th>range</th>
<th>further properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tilde{E})</td>
<td>(\tilde{U}^HBAU)</td>
<td>({0})</td>
<td>(\mathbb{C}^k)</td>
<td>assumed to be nonsingular</td>
</tr>
<tr>
<td>(\tilde{M})</td>
<td>(UE^{-1}\tilde{U}^H)</td>
<td>(\tilde{U}\perp)</td>
<td>(\mathbb{C})</td>
<td>rank (M = k)</td>
</tr>
<tr>
<td>(\tilde{P})</td>
<td>(I - AMB)</td>
<td>(\tilde{Z})</td>
<td>((B^H\tilde{U})\perp)</td>
<td>(P^2 = P)</td>
</tr>
<tr>
<td>(\tilde{Q})</td>
<td>(I - MBA)</td>
<td>(\tilde{U})</td>
<td>((A^HB^H\tilde{U})\perp)</td>
<td>(Q^2 = Q)</td>
</tr>
<tr>
<td>(\tilde{A})</td>
<td>(PA)</td>
<td>(\tilde{U})</td>
<td>((B^H\tilde{U})\perp)</td>
<td>(\tilde{A} = PA = AQ = PAQ)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>definition</th>
<th>null space</th>
<th>range</th>
<th>further properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{E})</td>
<td>(U^HB^HA^H\hat{U})</td>
<td>({0})</td>
<td>(\mathbb{C}^k)</td>
<td>assumed to be nonsingular</td>
</tr>
<tr>
<td>(\hat{M})</td>
<td>(\tilde{U}^H\tilde{U}^H)</td>
<td>(\hat{U}\perp)</td>
<td>(\mathbb{C})</td>
<td>rank (\tilde{M} = k)</td>
</tr>
<tr>
<td>(\hat{P})</td>
<td>(I - A^HM^B^H)</td>
<td>(\hat{Z})</td>
<td>((BU)\perp)</td>
<td>(\hat{P}^2 = \hat{P})</td>
</tr>
<tr>
<td>(\hat{Q})</td>
<td>(I - M^B^HA^H)</td>
<td>(\hat{U})</td>
<td>((ABU)\perp)</td>
<td>(\hat{Q}^2 = \hat{Q})</td>
</tr>
<tr>
<td>(\hat{A})</td>
<td>(\hat{P}A^H)</td>
<td>(\hat{U})</td>
<td>((BU)\perp)</td>
<td>(\hat{A} = \hat{P}A^H = A^H\hat{Q} = \hat{P}A^H\hat{Q})</td>
</tr>
</tbody>
</table>

Krylov spaces used:

\[
\hat{\mathcal{K}}_n \equiv \mathcal{K}_n(\hat{A}, \hat{r}_0), \quad \hat{\mathcal{L}}_n \equiv \mathcal{K}_n(\hat{A}, \hat{r}_0).
\]
**Theorem**

For $n \geq 1$ the two sets of conditions

\[
x_n \in x_0 + \hat{K}_n + U, \quad r_n \perp B^H(\hat{L}_n + \tilde{U}),
\]

\[
\tilde{x}_n \in \tilde{x}_0 + \hat{L}_n + \tilde{U}, \quad \tilde{r}_n \perp B(\hat{K}_n + U)
\]

and

\[
\hat{x}_n \in x_0 + \hat{K}_n, \quad \hat{r}_n \perp B^H\hat{L}_n,
\]

\[
\tilde{x}_n \in \tilde{x}_0 + \hat{L}_n, \quad \tilde{r}_n \perp B\hat{K}_n
\]

are equivalent in the sense that

\[
x_n = Q\hat{x}_n + MBb \quad \text{and} \quad r_n = \hat{r}_n, \quad (7)
\]

\[
\tilde{x}_n = \tilde{Q}\hat{x}_n + M^HB^H\tilde{b} \quad \text{and} \quad \tilde{r}_n = \tilde{r}_n. \quad (8)
\]
Deflated BiCG

For solving a single system $Ax = b$ with deflated BiCG the indirect approach requires to solve $\hat{A}\hat{x} = Pb$ such that

$$\hat{x}_n \in x_0 + \hat{K}_n, \quad \hat{r}_n \perp \hat{L}_n.$$  

This works with BiCG if $\tilde{A} = \tilde{A}^H$. Fortunately, we have:

**THEOREM**

If $B = I$, then $P^H = \tilde{Q}$ and $Q^H = \tilde{P}$, and therefore $\tilde{A}^H = \tilde{A}$.
Deflated BiCG

For solving a single system $\mathbf{Ax} = \mathbf{b}$ with deflated BiCG the indirect approach requires to solve $\hat{\mathbf{A}}\hat{\mathbf{x}} = \mathbf{Pb}$ such that

$$\hat{\mathbf{x}}_n \in \mathbf{x}_0 + \hat{\mathcal{K}}_n, \quad \hat{\mathbf{r}}_n \perp \hat{\mathcal{L}}_n.$$ 

This works with BiCG if $\hat{\mathbf{A}} = \hat{\mathbf{A}}^H$. Fortunately, we have:

**THEOREM**

If $\mathbf{B} = \mathbf{I}$, then $\mathbf{P}^H = \hat{\mathbf{Q}}$ and $\mathbf{Q}^H = \hat{\mathbf{P}}$, and therefore $\hat{\mathbf{A}}^H = \hat{\mathbf{A}}$.

It is not clear if one can obtain an equally efficient deflated BiCR (where $\mathbf{B} = \mathbf{A}$).
Deflated BICG

For solving a single system $Ax = b$ with deflated BICG the indirect approach requires to solve $\hat{A}\hat{x} = \hat{P}b$ such that

$$\hat{x}_n \in x_0 + \hat{K}_n, \quad \hat{r}_n \perp \hat{L}_n.$$ 

This works with BICG if $\tilde{A} = \tilde{A}^H$. Fortunately, we have:

**Theorem**

If $B = I$, then $P^H = \tilde{Q}$ and $Q^H = \tilde{P}$, and therefore $A^H = \tilde{A}$.  

It is not clear if one can obtain an equally efficient deflated BICR (where $B = A$).

But BICGSTAB, IDR(s), etc. can be done.
Conclusions

- We have set up a framework for deflated Krylov space solvers based on a Galerkin condition.
- We are developing an analogue framework for those based on a Petrov-Galerkin condition.
- By our indirect approach we obtain a deflated BiCG with coupled two-term recurrences (or three-term recurrences) which reduces to deflated CG when applied to an Hpd problem:
  - If $B = I$, BiCG applied to the projected problem $\hat{A}\hat{x} = P\hat{b}$ yields $\hat{x}_n$, $\hat{r}_n$, and shadow residuals $\tilde{r}_n$.
  - The transformation (7) yields $x_n$. 
References

