TRINITY COLLEGE DUBLIN

UPPSALA UNIVERSITET

Deflation techniques - historical development and advances

Maya Neytcheva

Department of Information Technology
Uppsala University

Plan of the lecture

- Introduction: a bit on Krylov methods for exactly singular and nearly singular systems
- Speed up the convergence of iterative methods by eliminating small eigenvalues
  - The Bordering method
  - Bordering as preconditioning
  - Deflation
- Some theoretical results
- Computing eigenvectors, given we know the eigenvalues
- A few numerical tests
- Conclusions, open questions

Notations

Given a matrix $A(n, n)$;
Eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$
Eigenvectors: $\boldsymbol{v}^{(1)}, \boldsymbol{v}^{(2)}, \cdots, \boldsymbol{v}^{(n)}$
$V_p = [\boldsymbol{v}^{(1)}, \boldsymbol{v}^{(2)}, \cdots, \boldsymbol{v}^{(p)}]$ - matrix of size (n, p)
$\mathcal{V} = \mathcal{V}_p = span(\boldsymbol{v}^{(1)}, \boldsymbol{v}^{(n)}, \cdots, \boldsymbol{v}^{(p)})$
$\mathcal{R}(B)$ - range of a matrix B
$\mathcal{N}(B)$ - nullspace of a matrix B

Assumed ordering of the eigenvalues:
A spd: $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$
A - general: $|\lambda_1| \leq |\lambda_2| \leq \cdots \leq |\lambda_n|$

Introduction: Deflation techniques

Deflation used in various contexts:

- in eigenvalue computations,
- when solving linear systems of equations via iterative soliton methods.

Deflation in eigenvalue computations

Y. Saad, *Numerical methods for large eigenvalue problems*, SIAM, 2011
Recall:

- Eigenvalue problem $A\boldsymbol{v} = \lambda \boldsymbol{v}$, $AV = V\Lambda$, right eigenvectors
- $U^T A = \Lambda U$, left eigenvectors. $U^T V = I$.

Assume that we have computed the largest modulus eigenvalue of A, λ_n and its eigenvector $\boldsymbol{v}^{(n)}$, $\|\boldsymbol{v}^{(n)}\| = 1$.
Question: How to compute next largest λ_{n-1} by 'hiding' (displacing, deflating) λ_{n-1} and not changing the rest of the spectrum?
Answer: By rank-one modification of A

$$\widetilde{A} = A - \alpha \boldsymbol{v}^{(n)} \boldsymbol{w}^H$$

$\boldsymbol{w}$ arbitrary, such that $\boldsymbol{w}^H \boldsymbol{v}^{(n)} = 1$, α - some shift.

Deflation in eigenvalue computations

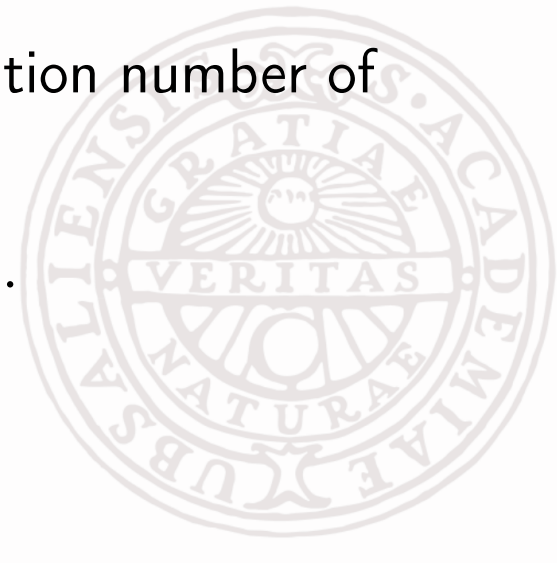
Theorem:(Wieland) $A_1 = A - \alpha \boldsymbol{v}^{(n)} \boldsymbol{w}^H$: the spectrum of A_1 is

$$\{\lambda_1, \cdots, \lambda_{n-1}, \lambda_n - \alpha\}.$$

How to choose $\boldsymbol{w}$?
Possible goal: $\boldsymbol{w}$ such that it minimizes the condition number of λ_{n-1}:

$$\varkappa(\lambda_{n-1}) = \frac{\|\widetilde{\boldsymbol{v}}^{(n-1)}\|_2 \|\widetilde{\boldsymbol{u}}^{(n-1)}\|_2}{((\widetilde{\boldsymbol{u}}^{(n-1)})^T \widetilde{\boldsymbol{v}}^{(n-1)})}.$$

How to choose α?
How to deflate several eigenvectors?
(important within Lanczos and Arnoldi methods)



Deflation and iterative solvers

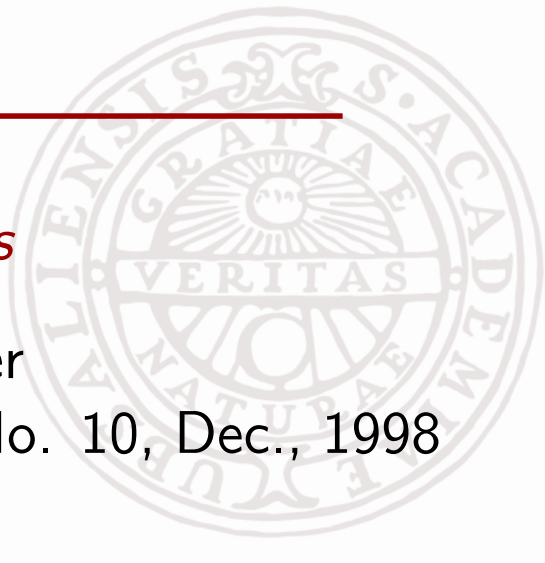
Iterative solution methods

Why Krylov subspace methods are so much used?

Since 1931.

The Idea Behind Krylov Methods

Ilse C. F. Ipsen and Carl D. Meyer
The American Mathematical Monthly, Vol. 105, No. 10, Dec., 1998



Properties of the Krylov subspaces

$\mathcal{K}^m(A, \boldsymbol{v}) = span\{\boldsymbol{v}, A\boldsymbol{v}, \cdots, A^{m-1}\boldsymbol{v}\}$
The dimension of $\mathcal{K}^m$ increases with each iteration.
Led d be the degree of the minimal polynomial of A.

- $\mathcal{K}^d$ is invariant under A, thus, $\mathcal{K}^m = \mathcal{K}^d$ for $m > d$:

$$dim(\mathcal{K}^m) = \min(m, d)$$

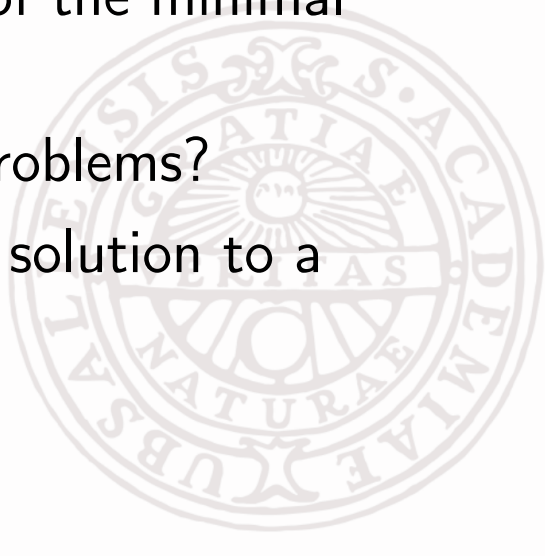
Summary:

Q1 Why Krylov methods make sense, and why it is natural to represent a solution to a linear system as a member of a Krylov space?

A1 The solution to a nonsingular linear system $Ax = b$ lies in a Krylov space whose dimension is the degree of the minimal polynomial of A.

Q2 Why Krylov methods work fine for singular problems?

A2 Under certain conditions there exist a Krylov solution to a singular problem.



Assume that A is nonsingular

Idea: express A^{-1} in terms of powers of A.

A1 The minimal polynomial of A, $p_d(t)$ of degree d, is the unique monic polynomial of minimal degree, for which

$$p_d(A) = 0.$$

It has the form

$$p_d(t) = \prod_{j=1}^{d} (t - \lambda_j)^{m_j},$$

where

- $\lambda_1, \cdots, \lambda_d$ are distinct eigenvalues of A,
- $m_1, \cdots, m_d$ are the corresponding indeces of λ_j (the sizes of the largest Jordan block, associated with λ_j).

Express A^{-1} in terms of powers of A.

$$p_d(t) = \prod_{j=1}^{d} (t - \lambda_j)^{m_j} = \sum_{s=0}^{m} \alpha_s t^s,$$

where $m = \sum_{j=1}^{d} m_j$.

Note that, since we have assumed that A is nonsingular, the coefficient α_0, $\alpha_0 = \prod_{j=1}^{d} (-\lambda_j)^{m_j}$ is not zero.

Express A^{-1} in terms of powers of A.

$$p_d(A) = \alpha_0 I_n + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_d A^d = 0,\ \alpha_0 \neq 0$$

Then $A^{-1} p_d(A) = 0$, thus,

$$A^{-1} = \frac{1}{\alpha_0} \sum_{j=0}^{d-1} \alpha_{j+1} A^j$$

However, $\boldsymbol{x} = A^{-1}\boldsymbol{b}$!

If the minimal polynomial of A ($A^{-1}\exists$) has degree d, then $\boldsymbol{x} = A^{-1}\boldsymbol{b} \in \mathcal{K}^d(A, \boldsymbol{b})$.
($\boldsymbol{b} = \boldsymbol{r}^{(0)}$ for $\boldsymbol{x}^{(0)} = \boldsymbol{0}$)

What happens if A^{-1} does not exist?

A2 When the matrix is singular, Krylov methods can fail.

Even if the linear system does have a solution, it may not lie in a Krylov space.

However, we can describe a class of right-hand sides for which a solution lies in a Krylov space. As it happens, there is only a single solution that lies in a Krylov space, and it can be obtained from the so-called Drazin inverse.

Assume A^{-1} does not exist

Apply the following trick: Decompose the space $\mathbb{C}^n = \mathcal{R}(A^\ell) \oplus \mathcal{N}(A^\ell)$, where ℓ is the index of the zero eigenvalue of $A \in \mathbb{C}^{n\times n}$. Then, there exists Q nonsingular,

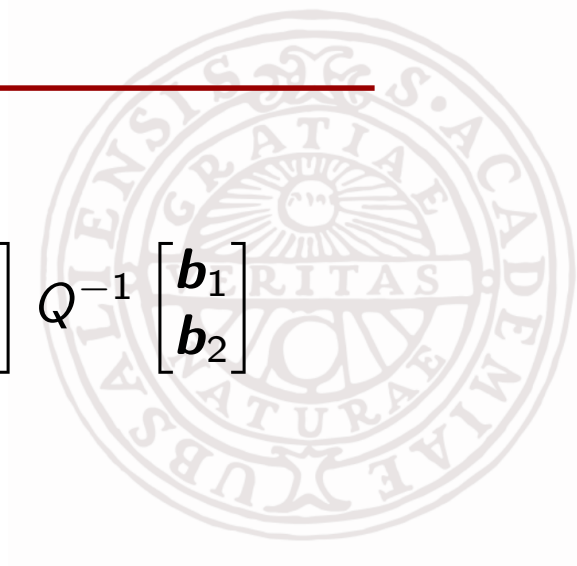
$$A = Q \begin{bmatrix} R & 0 \\ 0 & N \end{bmatrix} Q^{-1},$$

where R is nonsingular and N is nilpotent of index ℓ.

Suppose now that $A\boldsymbol{x} = \boldsymbol{b}$ has a Krylov solution

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} = \sum_{j=1}^{p} \alpha_j A^j \boldsymbol{b} = \sum_{j=0}^{p} \alpha_j Q \begin{bmatrix} R^j & 0 \\ 0 & N^j \end{bmatrix} Q^{-1} \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix}$$

Set $\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} = Q^{-1} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}$, $\begin{bmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \end{bmatrix} = Q^{-1} \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix}$.



Assume A^{-1} does not exist

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \sum_{j=1}^{p} \alpha_j A^j b = \sum_{j=0}^{p} \alpha_j Q \begin{bmatrix} R^j & 0 \\ 0 & N^j \end{bmatrix} Q^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Set $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = Q^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = Q^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$. Then

$$y_1 = \sum_{j=0}^{p} \alpha_j R^j z_1 \quad \text{and} \quad y_2 = \sum_{j=0}^{p} \alpha_j N^j z_2.$$

From $Ax = b$ we have that $Ny_2 = z_2$, so $\sum_{j=0}^{p} \alpha_j N^{j+1} z_2 = z_2$ and

$$(I - \sum_{j=0}^{p} \alpha_j N^{j+1}) z_2 = 0$$

The matrix in parentheses is nonsingular, thus $b_2 = 0$.
In other words, the existence of a Krylov solution requires that $b \in \mathcal{R}(A^\ell)$. The converse statement is also true.

Assume A^{-1} does not exist

Theorem

A square linear system $A\boldsymbol{x} = \boldsymbol{b}$ has a Krylov solution if and only if $\boldsymbol{b} \in \mathcal{R}(A^{\ell})$, where ℓ is the index of the zero eigenvalue of A.

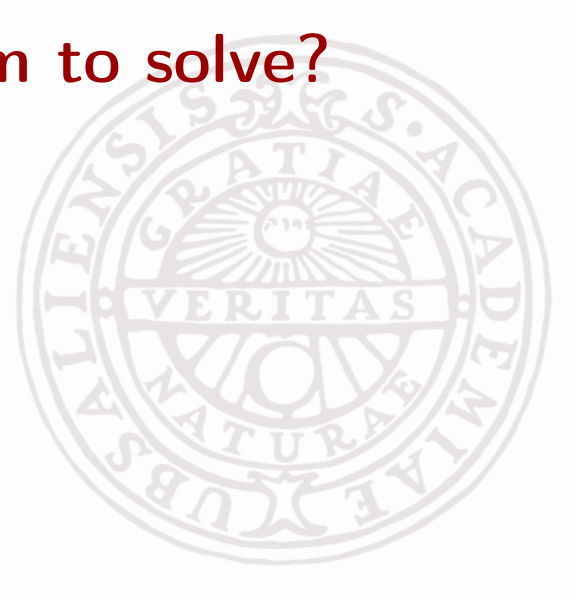
The solution is

$$\boldsymbol{x} = Q \begin{bmatrix} R^{-1}\boldsymbol{z}_1 \\ \boldsymbol{0} \end{bmatrix}.$$

Provided that $\boldsymbol{b}$ is confined to the "nonsingular part of A", the singular system has a Krylov solution.

Furthermore, the dimension of the Klylov subspace is smaller by ℓ - the index of the zero eigenvalue of A.

What if we have a nearly singular system to solve?



Troubles when solving nearly singular systems

Consider the problem of solving the system

$$A\boldsymbol{x} = \boldsymbol{b}$$

where $A(n \times n)$ is symmetric positive definite.
Let

$$\{\lambda_i, \boldsymbol{v}^{(i)}\},\ i = 1, 2, \cdots, n$$

be its eigensolutions and $\lambda_1 < \lambda_2 < \cdots < \lambda_n$.
If the vector $\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{v}^{(i)}$ is the solution vector, its components are found to be

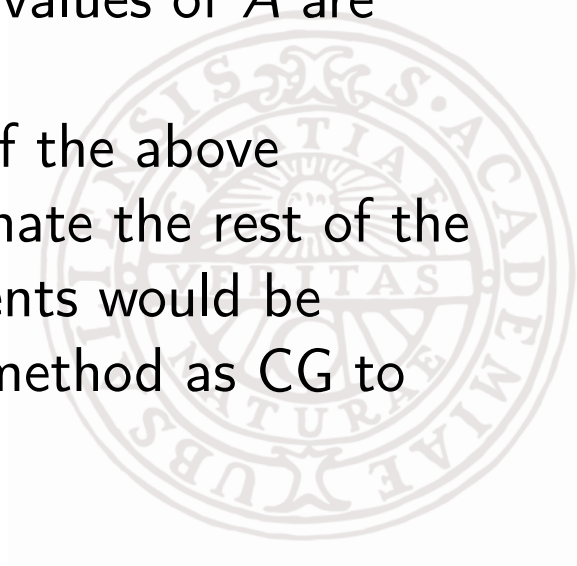
$$x_i = \frac{\boldsymbol{b}^T \boldsymbol{v}^{(i)}}{\lambda_i},\ i = 1, 2, \cdots, n.$$

Trouble makers: the very small eigenvalues...

Let $\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{v}^{(i)}$ be the solution of $A\boldsymbol{x} = \boldsymbol{b}$,

$$x_i = \frac{\boldsymbol{b}^T \boldsymbol{v}^{(i)}}{\lambda_i}, \ i = 1, 2, \cdots, n.$$

Assume that A is nearly singular, i.e. some eigenvalues of A are close to zero.
Let λ_1 be such an eigenvalue. Clearly, because of the above relations, in general the component x_1 will dominate the rest of the components x_j, $j = 2, \cdots, n$ and these components would be approximated poorly when using some iterative method as CG to solve the system.

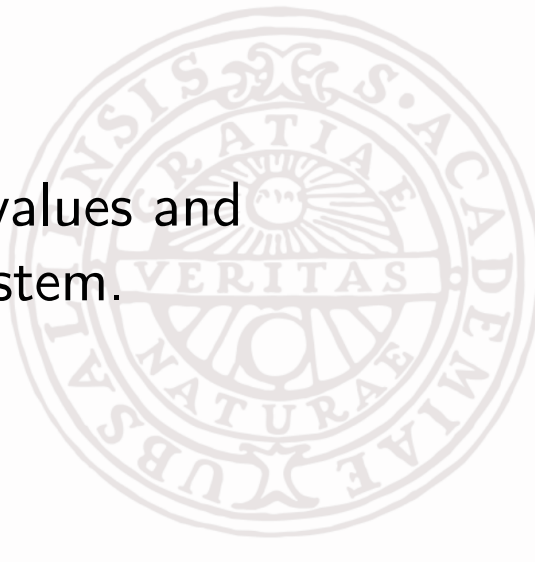


In contrast: CG for singular systems

Assume that A is exactly singular, i.e.,
$\lambda_1 = \cdots = \lambda_p = 0 < \lambda_{p+1} \leq \cdots \leq \lambda_n$ and $\boldsymbol{b}$ is in the range of A, then CG will converge to the solution of the system and its convergence will be improved, determined by

$$\frac{\lambda_n}{\lambda_{p+1}}.$$

Starting point to try to deflate the smallest eigenvalues and improve the conditioning of the resulting linear system.



The Bordering method

Source: D.K. Fadeev, V.N. Fadeeva, *Computational Methods of Linear Algebra*, Freeman, San Francisco, 1963

The idea: Let λ_1 close to zero and fairly well separated from λ_2. Let $\boldsymbol{w} \in \boldsymbol{R}^n$ be a vector such that $A\boldsymbol{w} \geq \boldsymbol{0}$, $A\boldsymbol{w} \neq \boldsymbol{0}$, $\boldsymbol{w}^T\boldsymbol{w} \neq \boldsymbol{0}$ and $\|\boldsymbol{w}\|_2 = 1$.
The system A is bordered with a row and a column to make a new system $\widetilde{A}(n+1 \times n+1)$ exactly singular,

$$\widetilde{A} = \begin{bmatrix} A & -\alpha A\boldsymbol{w} \\ -\alpha \boldsymbol{w}^T A & \alpha^2 \boldsymbol{w}^T A\boldsymbol{w} \end{bmatrix},$$

for some constant $\alpha > 0$.
Obviously, $\widetilde{A}$ is singular for $\widetilde{\boldsymbol{w}} = \begin{bmatrix} \alpha\boldsymbol{w} \\ 1 \end{bmatrix}$.

The Bordering method

Now we solve (straightforwardly using CG)

$$\widetilde{A}\widetilde{\boldsymbol{x}} = \widetilde{\boldsymbol{b}},$$

where $\widetilde{\boldsymbol{x}} = \begin{bmatrix} \boldsymbol{x} \\ \xi \end{bmatrix}$ and $\widetilde{\boldsymbol{b}} = \begin{bmatrix} \boldsymbol{b} \\ -\alpha \boldsymbol{w}^T \boldsymbol{b} \end{bmatrix}$.

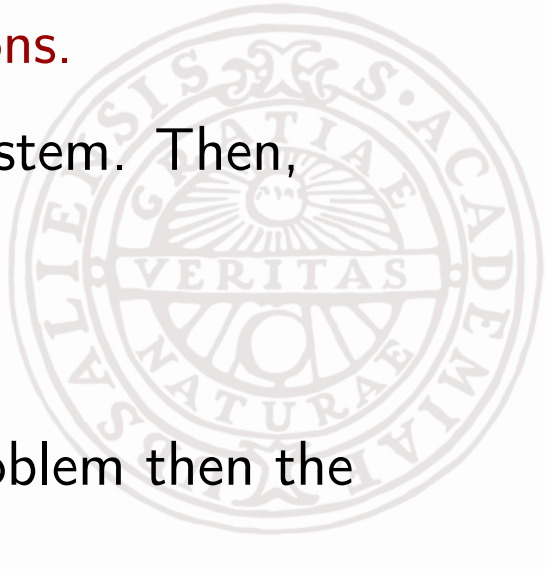
The augmented system has infinitely many solutions.

Let $\widetilde{\boldsymbol{x}}^* = \begin{bmatrix} \boldsymbol{x}^* \\ \xi^* \end{bmatrix}$ be a solution of the augmented system. Then,

$$\widehat{\boldsymbol{x}} = \boldsymbol{x}^* - \alpha \xi^* \boldsymbol{w}$$

is the solution of $A\boldsymbol{x} = \boldsymbol{b}$.

Analogously, if $\widehat{\boldsymbol{x}}$ is the solution of the original problem then the vector $\begin{bmatrix} \widehat{\boldsymbol{x}} \\ 0 \end{bmatrix}$ is a solution of the augmented one.



Bordering with special vectors

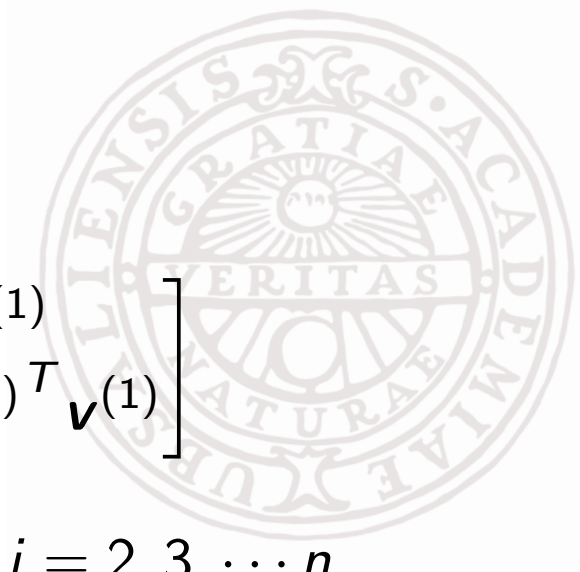
Choosing properly the vector $\boldsymbol{w}$ and the factor α we can *improve the condition number of* $\widetilde{A}$.

Let $\boldsymbol{w} = \boldsymbol{v}^{(1)}$. Then,for the eigenpairs $\{\widetilde{\lambda}_k, \widetilde{\boldsymbol{v}}^{(k)}\}, k = 0, 1, \cdots, n$ of $\widetilde{A}$ we have

$$\widetilde{\lambda}_0 = 0 \qquad \widetilde{\boldsymbol{v}}^{(0)} = \begin{bmatrix} \alpha \boldsymbol{v}^{(1)} \\ 1 \end{bmatrix}$$

$$\widetilde{\lambda}_1 = \lambda_1 (1 + \alpha^2 \boldsymbol{v}^{(1)\,T} \boldsymbol{v}^{(1)}) \qquad \widetilde{\boldsymbol{v}}^{(1)} = \begin{bmatrix} \boldsymbol{v}^{(1)} \\ -\alpha \boldsymbol{v}^{(1)\,T} \boldsymbol{v}^{(1)} \end{bmatrix}$$

$$\widetilde{\lambda}_j = \lambda_j \qquad \widetilde{\boldsymbol{v}}^{(j)} = \begin{bmatrix} \boldsymbol{v}^{(j)} \\ 0 \end{bmatrix}, \; j = 2, 3, \cdots n.$$



Bordering with special vectors, cont.

Thus, choosing the constant α so that

$$\lambda_2 < \lambda_1(1 + \alpha^2 \boldsymbol{v}^{(1)\,T} \boldsymbol{v}^{(1)}) < \lambda_n$$

we get

$$\varkappa(\widetilde{A}) = \frac{\lambda_n}{\lambda_2} < \frac{\lambda_n}{\lambda_1} = \varkappa(A).$$

Bordering with special vectors, cont.

If several nearly zero eigenvalues in the spectrum of A – border it with the corresponding number of columns and rows!

Theorem Let A be of order $n \times n$ and W_p of order $n \times p$ where $p < n$. Consider the augmented system

$$\widetilde{A} = \begin{bmatrix} A & -AW_p \\ -W_p^T A & W_p^T AW_p \end{bmatrix}.$$

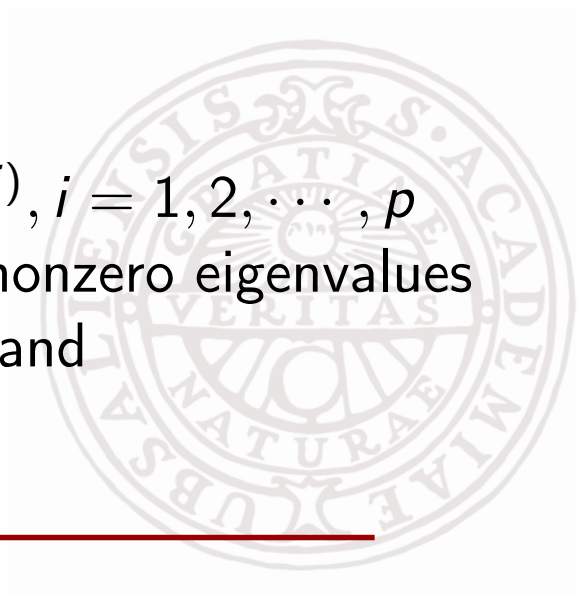
Bordering with special vectors, cont.

Then

a) $\widetilde{A}$ has p zero eigenvalues. There are n eigenvalues $\widetilde{\lambda}_i$ which are equal to those of $(I + W_p W_p^T)A$.

b) If A is spd, then to every eigenvalue λ_i of A there exists $\widetilde{\lambda}_i$ such that $\widetilde{\lambda}_i \geq \lambda_i$.

c) If A is nonsingular and symmetric and $W_p \equiv V_p = [\alpha_1 \boldsymbol{v}^{(1)}, \cdots, \alpha_p \boldsymbol{v}^{(p)}]$, where $\boldsymbol{v}^{(i)}, i = 1, 2, \cdots, p$ are normalized eigenvectors of A, then the nonzero eigenvalues of $\widetilde{A}$ equal $\widetilde{\lambda}_i = (1 + \alpha_i^2)\lambda_i, i = 1, 2, \cdots, p$ and $\widetilde{\lambda}_i = \lambda_i, i = p + 1, \cdots, n$.



Proof: $\widetilde{A}$ can be factorized as $\widetilde{A} = U^T Z U$,

$$\widetilde{A} = \begin{bmatrix} A & -AW_p \\ -W_p^T A & W_p^T A W_p \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -W_p^T & I_p \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I_n & -W_p \\ 0 & I_p \end{bmatrix},$$

so $\widetilde{\widetilde{A}}$, similar to $U\widetilde{A}U^{-1}$, has the form

$$\widetilde{\widetilde{A}} = \begin{bmatrix} I_n & -W_p \\ 0 & I_p \end{bmatrix} \begin{bmatrix} I_n & 0 \\ -W_p^T & I_p \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} (I_n + W_p W_p^T)A & 0 \\ -W_p^T A & 0 \end{bmatrix}.$$

The above shows part a); b) follows directly from a).
If $A\boldsymbol{v}^{(i)} = \lambda_i \boldsymbol{v}^{(i)}, i = 1, 2, \cdots, n$ then

$$(I_n + V_p V_p^T)A\boldsymbol{v}^{(i)} = \lambda_i (I_n + V_p V_p^T)\boldsymbol{v}^{(i)} = \begin{cases} (1+\alpha_i^2)\lambda_i \boldsymbol{v}^{(i)}, & i = 1, \cdots, p \\ \lambda_i \boldsymbol{v}^{(i)}, & i = p+1, \cdots, n \end{cases}$$

where use the orthogonality of the eigenvectors (part c). ∎

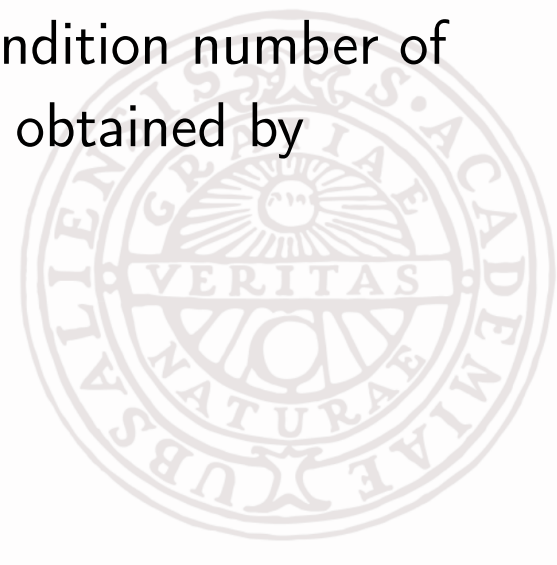
Choose $\alpha_i = \sqrt{\lambda_n/\lambda_i - 1}$.

Avoiding the solution of the augmented system

From the theorem it follows that we may also use $I + W_p W_p^T$ as a *preconditioner* to A. This will have the same convergence rate as CG on the augmented system $\widetilde{A}$ but has the advantage that we work with a smaller-sized system.

Theorem For $W_p = V_p$, the smallest effective condition number of $\widetilde{A}$, i.e., $\lambda_{max}((I + V_p^T V_p)A)/\lambda_{min}((I + V_p^T V_p)A)$ obtained by bordering with p vectors is

$$\varkappa_{eff}(A) = \lambda_n/\lambda_{p+1}.$$



Proof: Let $\Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$ be a diagonal matrix with the eigenvalues of A, V be orthonormal matrix with the eigenvectors of A as columns and $A = V\Lambda V^T$ be the spectral decomposition of A. Let also $V_p = VW_p$, where W is $n \times p$ matrix. Then

$$\begin{aligned} V^T(A + V_p V_p^T A)V &= V^T AV + V^T V_p V_p^T AV \\ &= \Lambda + V^T VW_p W_p^T V^T AV \\ &= \Lambda + W_p W_p^T \Lambda \qquad (\star) \\ &= \Lambda^{-\frac{1}{2}}(\Lambda + \Lambda^{\frac{1}{2}} W_p W_p^T \Lambda^{\frac{1}{2}})\Lambda^{\frac{1}{2}} \\ &= \Lambda^{-\frac{1}{2}}(\Lambda + \widetilde{W}_p \widetilde{W}_p{}^T)\Lambda^{\frac{1}{2}} \end{aligned}$$

where $\widetilde{W}_p = \Lambda^{\frac{1}{2}} W_p$ is $n \times p$. Since the column rank of W_p is less or equal p, then there exists a vector $\boldsymbol{z} = \{z_i\} \in \boldsymbol{R}^n$, such that

$$\widetilde{W}_p{}^T \boldsymbol{z} = \mathbf{0},\ z_i = 0,\ i = p+2, \cdots, n; \|\boldsymbol{z}\| = 1.$$

Proof, cont.: $\star$ shows a spectral equivalence relation $V^T(A + V_p V_p^T A)V = \Lambda^{-\frac{1}{2}}(\Lambda + \widetilde{W}_p \widetilde{W}_p{}^T)\Lambda^{\frac{1}{2}}$. Then

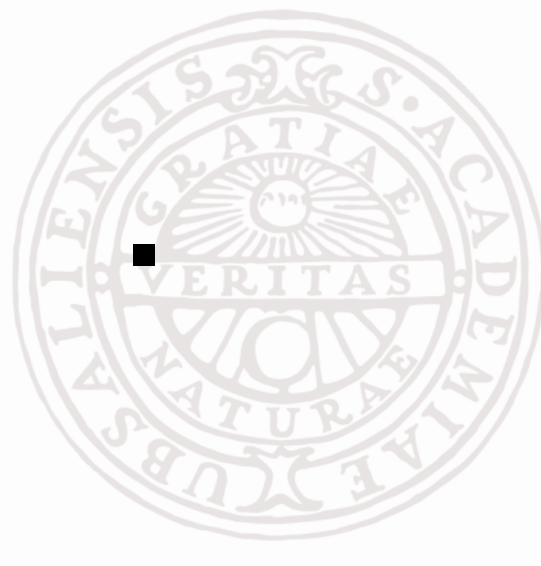
$$\lambda_{min}(A + V_p V_p^T A) = \lambda_{min}(\Lambda + \widetilde{W}_p \widetilde{W}_p{}^T) \leq (\Lambda z, z) + (\widetilde{W}_p \widetilde{W}_p{}^T z, z)$$

$$= \sum_{i=1}^{p+1} \lambda_i z_i^2 \leq \lambda_{p+1}.$$

We already have that

$$\lambda_{max}(A + V_p V_p^T A) \geq \lambda_n.$$

■



Observations:

$$\widetilde{A} = \begin{bmatrix} A & -AW_p \\ -W_p^T A & W_p^T AW_p \end{bmatrix}$$

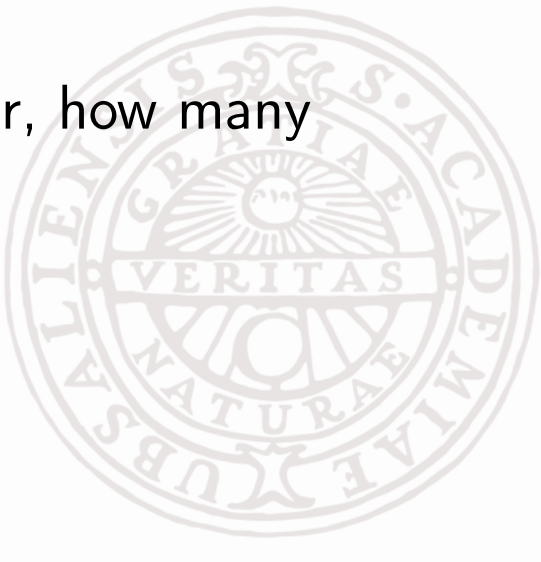
The Schur complement of $\widetilde{A}$ is

$$\begin{aligned} S_{\widetilde{A}} &= A - AW_p(W_p^T AW_p)^{-1} W_p^T A \\ &= A \underbrace{(I - W_p(W_p^T AW_p)^{-1} W_p^T A)}_{H} \\ &= AH \end{aligned}$$

$$\text{Also,} \quad AH = HA = H^T AH.$$

Relevant issues to clarify and discuss:

- How to incorporate the deflation idea into the CG, GMRES etc Krylov solvers?
- How to choose the vectors W_p?
- How sensitive is the corresponding solver to perturbations of W_p?
- Computation complexity of the deflated solver, how many vectors to use?



Vast literature!

Sources for this presentation:

1. Y. Saad, M. Yeung, J. Erhel, F. Guyomarch, A deflated version of the CG algorithm, SISC, 2000.
2. A. Padiy, O. Axelsson, B. Polman, Generalized augmented matrix preconditioning approach and its application to iterative solution of ill.conditioned algebraic systems, SIMAX, 2000.
3. K. Kahl, H. Rittich, The deflated CG method: convergence, perturbations and accuracy, LAA, 2017.
4. E. Carson, N. Knight, J. Demmel, An efficient deflation technique for the communication-avoiding CG, ETNA, 2014.
5. Many many other authors have contributed to the topic.

CG is based on the Lanczos algorithm:

Let: A - spd, build an orthonormal basis in the Krylov subspace:

Algorithm 1 Hermitian Lanczos algorithm

1: Choose $\boldsymbol{v}_1$ such that $\|\boldsymbol{v}_1\| = 1$, set $\beta_1 = 0$, $\boldsymbol{v}_0 = \boldsymbol{0}$.
2: **for** $j = 1, \cdots, p$ **do**
3: $\quad \boldsymbol{w}_j = A\boldsymbol{v}_j - \beta_j \boldsymbol{v}_{j-1}$
4: $\quad \alpha_j = \boldsymbol{w}_j^T \boldsymbol{v}_j$
5: $\quad \boldsymbol{w}_j = \boldsymbol{w}_j - \alpha_j \boldsymbol{v}_j$
6: $\quad \beta_{j+1} = \|\boldsymbol{w}_j\|_2$
7: $\quad \boldsymbol{v}_{j+1} = \boldsymbol{w}_j / \beta_{j+1}$
8: **end for**

Incorporate the deflation idea into the CG

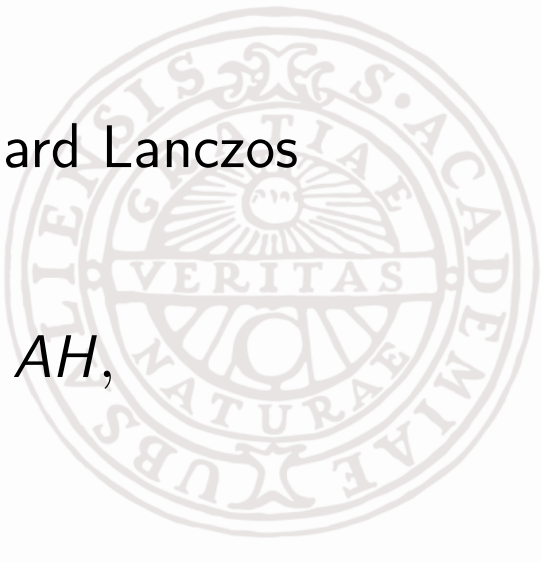
Let: A - spd. Choose $W_p = [\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_p]$ a set of linearly independent vectors.
The matrix $W^T AW$ is nonsingular.
Deflated Lanczos algorithm: build a sequence of vectors $\boldsymbol{v}_i$, s.t.

$$\boldsymbol{v}_{j+1} \perp span\{W, \boldsymbol{v}_1, \cdots, \boldsymbol{v}_j\}.$$

This turns out to be equivalent to apply the standard Lanczos procedure to the matrix

$$S_{\widetilde{A}} = A(I - W(W^T AW)^{-1} W^T A) = AH,$$

given a starting vector $\boldsymbol{v}_1$.



Deflated Lanczos algorithm:

1: Choose $W(n,p) = [\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_p]$.
2: Choose $\boldsymbol{v}_1$ such that $\|\boldsymbol{v}_1\| = 1$, $\boldsymbol{v}_1^T W = 0$. Set $\beta_1 \boldsymbol{v}_0 = \mathbf{0}$.
3: **for** $j = 1, \cdots, p$ **do**
4: Solve $W^T A W \boldsymbol{z}_j = W^T A \boldsymbol{v}_j$
5: $\boldsymbol{w}_j = A\boldsymbol{v}_j - AW\boldsymbol{z}_j$
6: $\alpha_j = \boldsymbol{v}_j^T \boldsymbol{w}_j$
7: $\boldsymbol{w}_j = \boldsymbol{w}_j - \beta_j \boldsymbol{v}_{j-1} - \alpha_j \boldsymbol{v}_j$
8: $\beta_{j+1} = \|\boldsymbol{w}_j\|$; if $\beta_{j+1} = 0$, exit
9: $\boldsymbol{v}_{j+1} = \boldsymbol{w}_j / \beta_{j+1}$
10: **end for**

Deflated CG algorithm:

1: Choose $\boldsymbol{x}^{(-1)}$, $\boldsymbol{r}^{(-1)} = A\boldsymbol{x}^{(-1)} - \boldsymbol{b}$, $W(n \times p)$ of full rank.
2: Compute and factorize $W^T AW$
3: $\boldsymbol{x}^{(0)} = \boldsymbol{x}^{(-1)} + W(W^T AW)^{-1} W^T \boldsymbol{r}^{(-1)}$, $\boldsymbol{r}^{(0)} = \boldsymbol{b} - A\boldsymbol{x}^{(0)}$.
4: Solve $W^T AW\boldsymbol{h} = W^T A\boldsymbol{r}^{(0)}$, $\boldsymbol{d}^{(0)} = \boldsymbol{r}^{(0)} - W\boldsymbol{h}$
5: **for** $k = 0, 1, \cdots$ until convergence **do**
6: $\quad \tau_k = \frac{(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)})}{(\boldsymbol{d}^{(k)}, A\boldsymbol{d}^{(k)})}$
7: $\quad \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \tau_k \boldsymbol{d}^{(k)}$
8: $\quad \boldsymbol{r}^{(k+1)} = \boldsymbol{r}^{(k)} + \tau_k A\boldsymbol{d}^{(k)}$
9: $\quad \beta_k = \frac{(\boldsymbol{r}^{(k+1)}, \boldsymbol{r}^{(k+1)})}{(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)})}$
10: $\quad$ Solve $W^T AW\boldsymbol{h} = W^T A\boldsymbol{r}^{(k+1)}$
11: $\quad \boldsymbol{d}^{(k+1)} = \boldsymbol{r}^{(k+1)} + \beta_k \boldsymbol{d}^{(k)} - W\boldsymbol{h}$
12: **end for** return $\boldsymbol{x}^{(k+1)}$

Deflated CG algorithm: concerns

1: Choose $\boldsymbol{x}^{(-1)}$, $\boldsymbol{r}^{(-1)} = A\boldsymbol{x}^{(-1)} - \boldsymbol{b}$, $W(n \times p)$ of full rank.

2: Compute and factorize $W^T AW$

3: $\boldsymbol{x}^{(0)} = \boldsymbol{x}^{(-1)} + W(W^T AW)^{-1} W^T \boldsymbol{r}^{(-1)}$, $\boldsymbol{r}^{(0)} = \boldsymbol{b} - A\boldsymbol{x}^{(0)}$.

4: Solve $W^T AW\boldsymbol{h} = W^T A\boldsymbol{r}^{(0)}$, $\boldsymbol{d}^{(0)} = \boldsymbol{r}^{(0)} - W\boldsymbol{h}$

5: **for** $k = 0, 1, \cdots$ until convergence **do**

6: $\quad \tau_k = \frac{(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)})}{(\boldsymbol{d}^{(k)}, A\boldsymbol{d}^{(k)})}$

7: $\quad \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \tau_k \boldsymbol{d}^{(k)}$

8: $\quad \boldsymbol{r}^{(k+1)} = \boldsymbol{r}^{(k)} + \tau_k A\boldsymbol{d}^{(k)}$

9: $\quad \beta_k = \frac{(\boldsymbol{r}^{(k+1)}, \boldsymbol{r}^{(k+1)})}{(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)})}$

10: $\quad$ Solve $W^T AW\boldsymbol{h} = W^T A\boldsymbol{r}^{(k+1)}$

11: $\quad \boldsymbol{d}^{(k+1)} = \boldsymbol{r}^{(k+1)} + \beta_k \boldsymbol{d}^{(k)} - W\boldsymbol{h}$

12: **end for** return $\boldsymbol{x}^{(k+1)}$

Preconditioned Deflated CG algorithm:

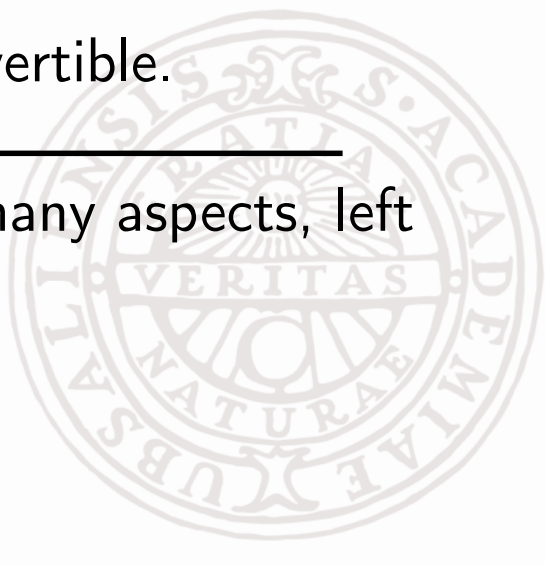
1: Choose $\boldsymbol{x}^{(-1)}$, $\boldsymbol{r}^{(-1)} = A\boldsymbol{x}^{(-1)} - \boldsymbol{b}$, $W(n \times p)$ of full rank.
2: Compute and factorize $W^T AW$
3: $\boldsymbol{x}^{(0)} = \boldsymbol{x}^{(-1)} + W(W^T AW)^{-1} W^T \boldsymbol{r}^{(-1)}$, $\boldsymbol{r}^{(0)} = \boldsymbol{b} - A\boldsymbol{x}^{(0)}$. Solve $P\boldsymbol{z}^{(0)} = \boldsymbol{r}^{(0)}$
4: Solve $W^T AW\boldsymbol{h} = W^T A\boldsymbol{z}^{(0)}$, $\boldsymbol{d}^{(0)} = \boldsymbol{r}^{(0)} - W\boldsymbol{h}$
5: **for** $k = 0, 1, \cdots$ until convergence **do**
6: $\quad \tau_k = \frac{(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)})}{(\boldsymbol{d}^{(k)}, A\boldsymbol{d}^{(k)})}$
7: $\quad \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \tau_k \boldsymbol{d}^{(k)}$
8: $\quad \boldsymbol{r}^{(k+1)} = \boldsymbol{r}^{(k)} + \tau_k A\boldsymbol{d}^{(k)}$
9: $\quad$ Solve $P\boldsymbol{z}^{(k+1)} = \boldsymbol{r}^{(k+1)}$
10: $\quad \beta_k = \frac{(\boldsymbol{r}^{(k+1)}, \boldsymbol{z}^{(k+1)})}{(\boldsymbol{r}^{(k)}, \boldsymbol{z}^{(k)})}$
11: $\quad$ Solve $W^T AW\boldsymbol{h} = W^T A\boldsymbol{z}^{(k+1)}$
12: $\quad \boldsymbol{d}^{(k+1)} = \boldsymbol{z}^{(k+1)} + \beta_k \boldsymbol{d}^{(k)} - W\boldsymbol{h}$
13: **end for** return $\boldsymbol{x}^{(k+1)}$

Breakdowns for the deflated Krylov subspace methods

CG: **No breakdowns!** The deflated CG computes the minimum A-norm solution over the space $\boldsymbol{x}^{(0)} + \mathcal{K}_{j,p}(A, W, \boldsymbol{r}^{(0)})$.

GMRES: **No breakdowns if** $\mathcal{N}(\widehat{H}A) \cap \mathcal{R}(\widehat{H}A) = \{0\}$,
where $\widehat{H} = I - AW(W^H A^H AW)^{-1} W^H A^H$,
$W^H A^H AW$ - Hermitian positive definite, thus, invertible.

Remark: Deflation and GMRES - there are very many aspects, left out of this presentation.



How to choose the deflation vectors?

Reduction of the condition number of A was shown if W consists of the exact eigenvectors of A, corresponding to the smallest p eigenvalues, thus $\mathcal{W}_p = \mathcal{V}_p$.

$$\|\boldsymbol{e}^{(j)}\| \leq 2 \left(\frac{\sqrt{\varkappa_{\mathit{eff}}} - 1}{\sqrt{\varkappa_{\mathit{eff}}} + 1} \right)^j \|\boldsymbol{e}^{(0)}\|, \quad j = 0, 1, \cdots$$

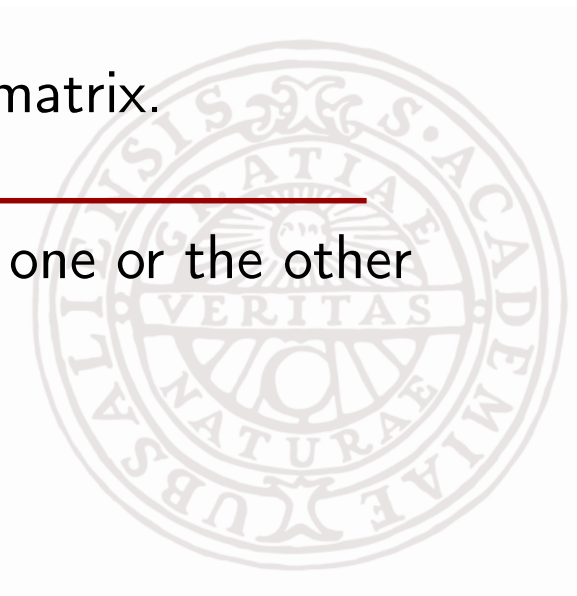
In practice these are not known.

How to choose the deflation vectors?

Strategies:

- Choose a subspace, $\mathcal{W}$, spanned by vectors $\boldsymbol{w}_j, j = 1, 2, \cdots, p$, either containing some computed approximations of the eigenvectors or vectors, based on additional knowledge of the problem.
- Approximate the matrix $V^T AV$ by another matrix.

Ask the question what will be the effect of using one or the other approach on the convergence of the deflated CG.

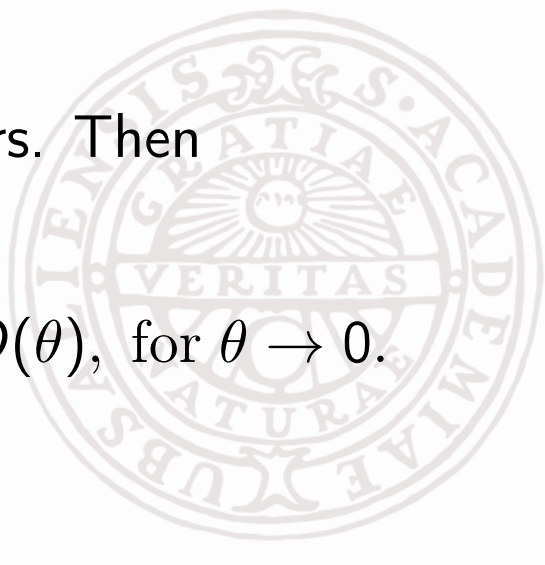


Strategy 1: Choose a subspace, different from V_p

Note: The difference between two subspaces is measured in terms of the largest principle angle θ between the subspaces.
If $\cos(\theta) = 0$, the spaces are orthogonal.
If $\sin(\theta)$ is small, we can expect that the subspaces are 'close' to each other.

Theorem: $\lambda(A) : \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.
$\mathcal{W}_p \neq \mathcal{V}_p$, e.g., contains approximated eigenvectors. Then

$$\varkappa_{\mathit{eff}} \leq \left(\sqrt{\frac{\lambda_n}{\lambda_{p+1}}} + \sqrt{\frac{\lambda_n}{\lambda_1}} \sin(\theta) \right)^2 = \frac{\lambda_n}{\lambda_{p+1}} + O(\theta), \text{ for } \theta \to 0.$$



Strategy 1: Choose a subspace, different from V_p

Consider a preconditioner $B = I + \sigma W(W^T AW)^{-1} W^T$,
$\sigma = \lambda_{max}(A)$,
$span(\boldsymbol{v}^{(1)}, \boldsymbol{v}^{(2)}, \cdots, \boldsymbol{v}^{(p)}) \subseteq \mathcal{R}(W)$. Then

$$\varkappa(BA) \leq \frac{2\lambda_{max}(A)}{\lambda_{p+1}}.$$

Strategy 2: Approximate $A_V = V^T AV$ / $A_W = W^T AW$

Replace $A_W = W^T AW$ by some other 'easier' matrix B_W:

Let $\widehat{B} = I + \widehat{\sigma} V B_W^{-1} V^T$, $\widehat{\sigma} = \frac{\lambda_{max}(A)}{\lambda_{max}(B_W^{-1} A_W)}$, $V \subseteq W$, then

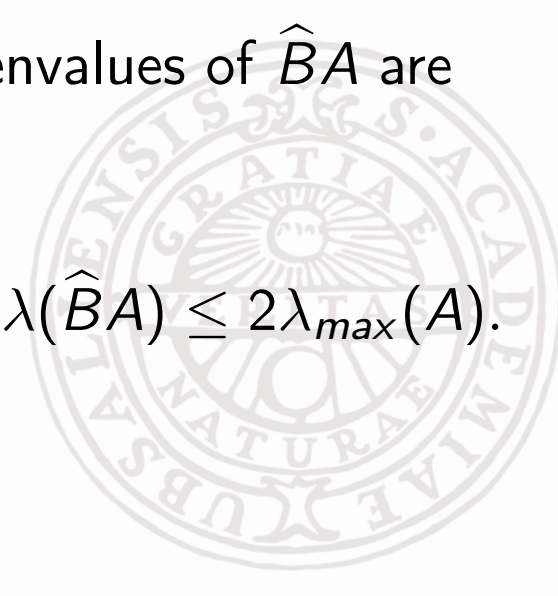
$$\min \left\{ \frac{\lambda_{max}}{\varkappa(\widehat{B}_W^{-1} A_W)} + \lambda_1, \lambda_{p+1} \right\} \leq \lambda(\widehat{B}A) \leq 2\lambda_{max}(A).$$

Theorem: Consider the preconditioner $\widehat{B} = I + \widehat{\sigma} W B_W^{-1} W^T$, $\widehat{\sigma} = \lambda_{max}(A)/\lambda_{max}(\widehat{B}_W^{-1} A_W)$, $\mathcal{V} = span\{\boldsymbol{v}^{(1)}, \cdots, \boldsymbol{v}^{(p)}\}$ and W is such that

$$\cos(\widetilde{\mathcal{W}}, \mathcal{V}) = \sup_{\substack{\boldsymbol{x} \in \widetilde{\mathcal{W}} \\ \boldsymbol{y} \in \mathcal{V}}} \frac{\boldsymbol{x}^T \boldsymbol{y}}{\|\boldsymbol{x}\| \, \|\boldsymbol{y}\|} \leq \gamma,$$

for $\widetilde{\mathcal{W}} = (\mathcal{R}(A^{1/2} W))^{\perp}$ and $\gamma < 1$, then the eigenvalues of $\widehat{B}A$ are bounded as

$$\max \left\{ \lambda_1, (1-\gamma) \min \left\{ \frac{\lambda_{max}(A)}{\varkappa(\widehat{B}_W^{-1} A_W)}, \lambda_{p+1} \right\} \right\} \leq \lambda(\widehat{B}A) \leq 2\lambda_{max}(A).$$

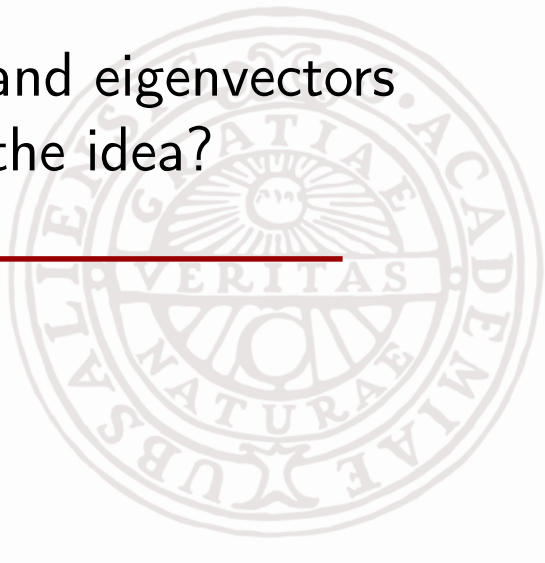


The difficulty: we do not know the eigenpairs

Issues to take care of:

- If W are based on approximate eigenvectors, how to approximate the small eigenpairs?
- Alternatively, how to find a subspace of vectors spanned by the eigenvectors?
- How accurate the approximated eigenvalues and eigenvectors should be in order to utilize the potential of the idea?

Can GLT help?



Deflated CG with the exact eigenvalues/vectors

1: Choose $\boldsymbol{x}^{(-1)}$, $\boldsymbol{r}^{(-1)} = A\boldsymbol{x}^{(-1)} - \boldsymbol{b}$, $W(n \times p)$ of full rank.

2: $W^T AW = \Lambda$ – diagonal matrix, nothing to factorize!

3: $\boldsymbol{x}^{(0)} = \boldsymbol{x}^{(-1)} + W\,(\Lambda)^{-1}\,W^T\boldsymbol{r}^{(-1)}$, $\boldsymbol{r}^{(0)} = \boldsymbol{b} - A\boldsymbol{x}^{(0)}$.

4: Solve $\Lambda \boldsymbol{h} = W^T A\boldsymbol{r}^{(0)}$, $\boldsymbol{d}^{(0)} = \boldsymbol{r}^{(0)} - W\boldsymbol{h}$

5: **for** $k = 0, 1, \cdots$ until convergence **do**

6: $\quad \tau_k = \frac{(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)})}{(\boldsymbol{d}^{(k)}, A\boldsymbol{d}^{(k)})}$

7: $\quad \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \tau_k \boldsymbol{d}^{(k)}$

8: $\quad \boldsymbol{r}^{(k+1)} = \boldsymbol{r}^{(k)} + \tau_k A\boldsymbol{d}^{(k)}$

9: $\quad \beta_k = \frac{(\boldsymbol{r}^{(k+1)}, \boldsymbol{r}^{(k+1)})}{(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)})}$

10: $\quad$ Solve $\Lambda \boldsymbol{h} = W^T A\boldsymbol{r}^{(k+1)}$

11: $\quad \boldsymbol{d}^{(k+1)} = \boldsymbol{r}^{(k+1)} + \beta_k \boldsymbol{d}^{(k)} - W\boldsymbol{h}$

12: **end for** return $\boldsymbol{x}^{(k+1)}$

GLT: example

$$-\varepsilon_1 u_{xx} - \varepsilon_2 u_{yy} = f$$

The symbol:

$$\begin{aligned} f_1(t_1) &= \varepsilon_1(2 - \cos(t_1)) \\ f_2(t_2) &= \varepsilon_2(2 - \cos(t_2)) \\ f(t_1, t_2) &= f_1(t_1) + f_2(t_2) \end{aligned}$$

Sampling: $h = \pi/(n+1)$

Eigenvectors:

In 1D: $\boldsymbol{v}_x^{(k)} = \sqrt{\frac{2}{n+1}} \sin(kt)$

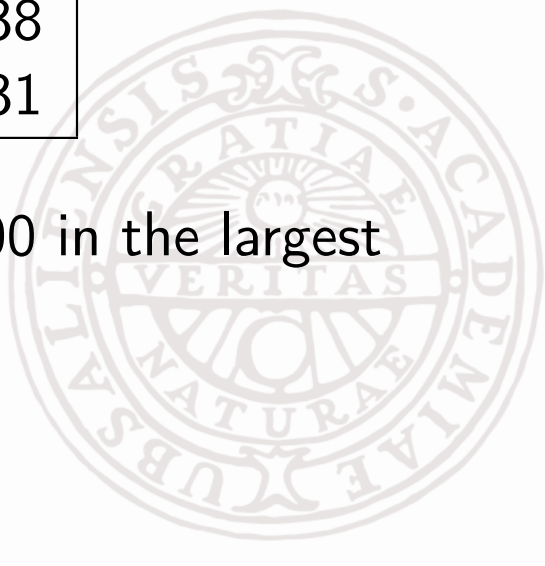
In 2D: $\boldsymbol{v} = \boldsymbol{v}_x \otimes \boldsymbol{v}_y$

GLT: $-\varepsilon_1 u_{xx} - \varepsilon_2 u_{yy} = f$

We know the symbol, the exact eigenvalues and the eigenvectors.

	Deflated CG		CG
Size	No deflated	iter.	iter
16384	1664	19	458
65536	6400	21	1038
262144	1000	166	2231

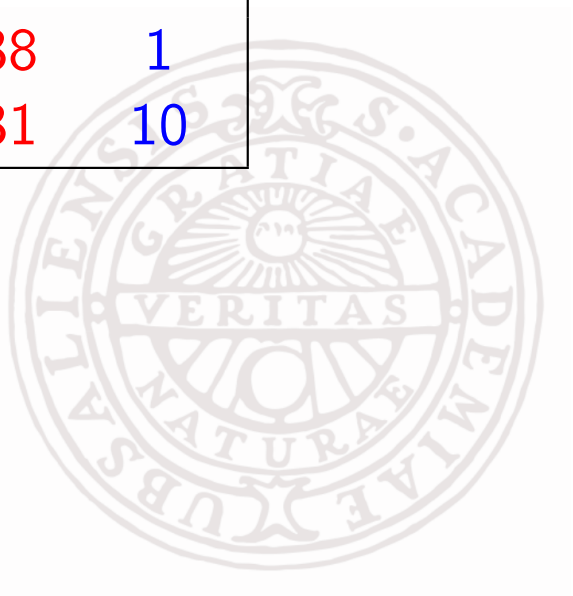
Deflate all eigenvalues, less than 0.1 (the first 1000 in the largest case).



GLT: $-\varepsilon_1 u_{xx} - \varepsilon_2 u_{yy} = f$

We know the symbol, the exact eigenvalues and the eigenvectors.

	Deflated CG			CG	
Size	No deflated	iter	time	iter	time
16384	1664	19	0.68	458	0.15
65536	6400	21	10	1038	1
262144	1000	166	54	2231	10



Some more numerical tests

Problem 1 (1D):

$$\begin{aligned} -\Delta^2 u &= f \qquad \text{in } \Omega = [0,1] \\ b.c. & \end{aligned}$$

The continuous biharmonic problem is discretized using cubic Hermite finite elements.

Numerical tests, biharmonic problem:

							CG		
n	Size	λ_{min}	p	λ_p	λ_{p+1}	λ_{max}	it 1	it 2	it 3
101	204	10^{-7}	99	0.4998	1.00	2.00	16	5	5
1001	2004	10^{-11}	999	0.5000	1.00	2.00	12	4	5
10001	20004	10^{-15}	9999	0.5000	1.00	2.00	8	4	5

- it 1: Number of iterations using no eigenvectors
- it 2: Number of iterations using p exact eigenvectors
- it 3: Number of iterations using p approximate eigenvectors, with an error $\mathcal{O}(10^{-7})$ ($\mathcal{O}(10^{-10}) \rightarrow$ 4 iterations)

Numerical tests: IgA

Problem 2: IgA (1D & 2D & 3D)

$$-\Delta u = f \qquad \text{in } \Omega = [0,1]^d, d = 1,2,3$$
$$b.c.$$

Laplace discretized with Isogeometric Analysis, polynomial degree p, and smoothness $k = p - 1 = 1$ (C^1)

$$\begin{aligned} 1D: &\quad nK_n\mathbf{u} = \mathbf{f} \\ 2D: &\quad n(K_n \otimes M_n + M_n \otimes K_n)\mathbf{u} = \mathbf{f} \\ 3D: &\quad ((K_n \otimes M_n) \otimes M_n) + ((M_n \otimes K_n) \otimes M_n) + ((M_n \otimes M_n) \otimes K_n) \end{aligned}$$

We know closed form formulae for eigenvalues and eigenvectors.
For $p > 2$ we can approximate the eigenvalues to arbitrary accuracy.

Numerical tests, IgA I

Problem 2: IgA (1D), cont.

$$\{K_n^{[2]}\}_n \sim_{glt,\lambda} f(\theta) = 1 - \frac{2}{3}\cos(\theta) - \frac{1}{3}\cos(2\theta)$$

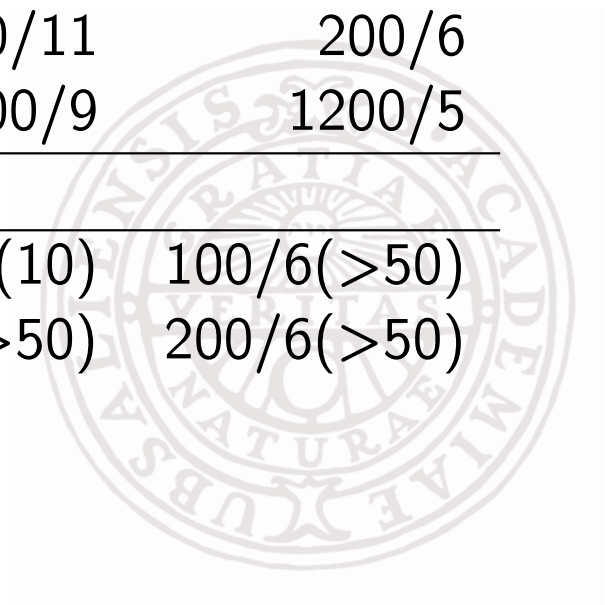
$$\lambda_j(K_n^{[2]}) = f(\theta_{j,n}) \qquad \theta_{j,n} = \frac{j\pi}{n}$$

$$x_{i,j} = \sqrt{\frac{2}{n}}\left[k_j \sin\left(\frac{(2i-1)j\pi}{2n}\right)\right]_{i=1}^{n} \qquad k_j = \begin{cases} 1/\sqrt{2} & \text{if } j = n \\ 1 & \text{otherwise} \end{cases}$$

Extendable to higher dimensions using Kronecker products.

Numerical tests, IgA II

1D	No. deflation vectors / CG iterations			
Size	No deflation	$1/4[< 0.5]$	$1/2[< 0.5]$	$[< 0.5]$
		$p = 2$		
500	0/310	25/17	50/10	100/6
1000	0/618	50/17	100/11	200/6
5000	0/2121	300/14	600/9	1200/5
		$p = 3$		
500	0/271	25/14(15)	50/10(10)	100/6(>50)
1000	0/536	50/16 (>50)	100/10(>50)	200/6(>50)



Numerical tests, IgA III

2D	No. deflation vectors / CG iterations			
Size	No deflation	$1/4[< 0.5]$	$1/2[< 0.5]$	$[< 0.5]$
		$p = 2$		
50^2	0/37	25/17	50/10	100/6
100^2	0/70	50/14	100/10	200/8

Exact eigenvalues and approximate eigenvectors

Assume we know the eigenvalues exactly!

But what about the eigenvectors?!

Two cases:

- We know those too (easy). For example IgA $p = 1, 2, 3$, some FDM.
- We know only an approximation of those (more difficult). What can we do?!

Case A Given the exact eigenvalues and approximations of the eigenvectors, compute the exact eigenvectors.

Case B Use the exact eigenvalues and the approximated eigenvectors in the deflated Krylov iterative methods. (ongoing work)

Case C Exploit an expansion for eigenvectors too (ongoing work)
A. Böttcher, J. M. Bogoya, S. M. Grudsky, E. Maximenko
Asymptotic formulas for the eigenvalues and eigenvectors of Toeplitz matrices. Sbornik Mathematics 208(11) (2018)

Case A: Use a nonlinear framework

for $k = 1 : p$ **do**
  $\boldsymbol{v}$ approximation of $\boldsymbol{v}^{(k)}$
  $\boldsymbol{w} = \mathbf{0}$
  $J = A - \lambda_k I$
  while $\|\boldsymbol{v} - \boldsymbol{w}\| > 10^{-14}$ **do**
    $\boldsymbol{w} = \boldsymbol{v}$
    $\boldsymbol{r} = J\boldsymbol{w}$
    if $\|\boldsymbol{r}\| < 10^{-13}$ then stop
    Solve $J\boldsymbol{\delta_v} = \boldsymbol{r}$ using CG
    Set $\boldsymbol{v} = \boldsymbol{v} - \boldsymbol{\delta_v}$
  end while
  $W(:, k) = \boldsymbol{v}/\|\boldsymbol{v}\|$
end for

Case A, cont.:

If $A = A^T$, solve the Sylvester equation

$$AV - V\Lambda = \widetilde{V},$$

where $\widetilde{V}$ contains the eigenvector approximations; Λ - diagonal, contains the eigenvalues. Afterwards, orthogonalize V, update $\widetilde{V}$ and repeat. Converges cubically and behaves quite well in terms of global convergence (P.-A. Absil, R. Mahony, R. Sepulchre and P. Van Dooren, SIAM review paper).

If $A \neq A^T$ – stabilize the iteration, e.g., by adding a constraint:

$$AV - V\Lambda = \widetilde{V}, \qquad \widetilde{V}^T V = I$$

Solve possibly via a variant of Bartels-Stewart; local quadratic convergence expected, but the global convergence – less robust.

Correspondence with Daniel Kressner.

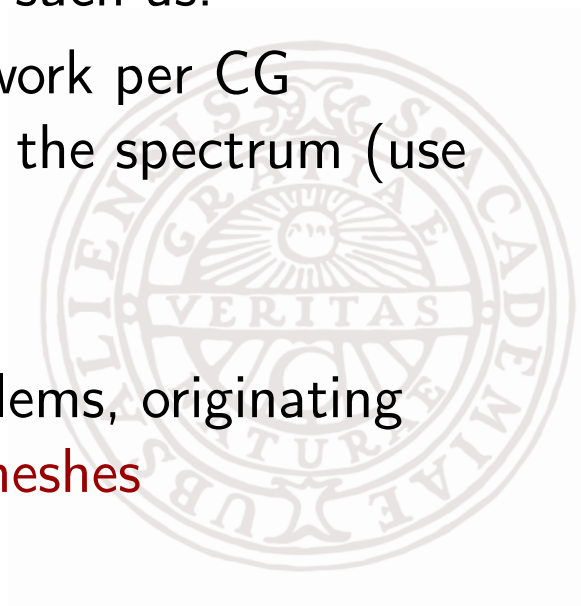
Conclusions, open problems

Deflation, powered with GLT may become an efficient building block in solving large scale problems, in particular on the modern HPC platforms.

There are many aspects to be further addressed, such as:

- How many vectors it is worthwhile to use (work per CG iteration increases). Problem-specific, study the spectrum (use the symbol)
- How to compute the eigenvectors efficiently
- Expand the technique to discrete PDE problems, originating from unstructured (triangular/tetrahedra) meshes
- How to automatically compute the symbol

Much to do!



Thank you for listening! Questions?