# Numerical Linear Algebra for Computational Science and Information Engineering

## Lecture 12
## Jacobi-Davidson method

Nicolas Venkovic
nicolas.venkovic@tum.de

Group of Computational Mathematics
School of Computation, Information and Technology
Technical University of Munich

Summer 2025

# Outline

# Motivation

- **Rayleigh-Ritz projections in Krylov subspaces** from Lanczos/Arnoldi procedures are very **effective to compute exterior eigenpairs**, provided the targeted eigenvalues are well-separated from the rest of the spectrum.
- In combination with a **shift-and-invert** spectral transformation, Rayleigh-Ritz projections in Krylov subspaces are also **efficient to compute interior eigenpairs** in the vicinity of a shift $\sigma$.

  A proper implementation of shift-and-invert transformations requires an evaluation of the mapping $x \mapsto (A - \sigma I_n)^{-1} x$ at each iteration.
- The **Jacobi-Davidson** method is of particular interest when one cannot afford to evaluate $x \mapsto (A - \sigma I_n)^{-1} x$ with sufficient precision.

  The Jacobi-Davidson method was proposed by Sleijpen and Van der Vorst (2000) on the basis of ideas from Jacobi (1845-46) and Davidson (1975).


Sleijpen, G. L., & Van der Vorst, H. A. (1996). A Jacobi–Davidson iteration method for linear eigenvalue problems. SIAM review, 42(2), 267-293.

Jacobi, C.G.J. (1845), Ueber eine neue Auflosungsart der bei der Methode der kleinsten Quadrate vorkommende linearen Gleichungen, Astronom Nachr, 297-306.

Jacobi, C. G. J. (1846). Über ein leichtes Verfahren die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen. J. Reine Angew. Math., 30, 51–94.

Davidson E. R. (1975). The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices. J. Comput. Phys., 17, 87–94.

# Davidson method

# Davidson method

- In the Davidson (1975) method, one is equipped with an **orthonormal basis** in the columns of $Q_k := [q_1, \ldots, q_k] \in \mathbb{F}^{n \times k}$.

  A **Rayleigh-Ritz projection** in range$(Q_k)$ is deployed where range$(Q_k)$ is ***not*** **a Krylov subspace**. We search for $(\lambda, \hat{y}) \in \mathbb{F} \times \mathbb{F}^k$ such that

$$B_k \hat{y} = \lambda \hat{y}$$

  where $B_k := Q_k^H A Q_k$. We then have a Rayleigh-Ritz pair $(\lambda, y)$ in which $y := Q_k \hat{y}$, and a residual given by $r := Ay - \lambda y$.
- The purpose of the Davidson method is to **improve the Rayleigh-Ritz vector** $y$ to decrease the residual norm $\|r\|_2$.

  For this, Davidson (1975) suggests to compute the **expansion vector** $t \in \mathbb{F}^n$ such that

$$(D_A - \lambda I_n) t = r$$

  where $D_A \in \mathbb{F}^{n \times n}$ is the diagonal matrix formed with the diagonal $A$.

Davidson E. R. (1975). The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices. J. Comput. Phys., 17, 87–94.

# Davidson method, cont'd$_1$

- The **expansion vector** $t$ is then **orthogonalized** against $q_1, \ldots, q_k$, **normalized** and **used to expand the search space**:

| |
|---|
| Solve for $t$ such that $(D_A - \lambda I_n)t = r$ |
| $t := \Pi^{(k)} t$ $\quad$ $//\Pi^{(k)}$ is a projector onto range$(Q_k)^\perp$ |
| $q_{k+1} := t/\|t\|_2$ |

  New Rayleigh-Ritz pairs can then be sought in range$(Q_{k+1})$.
- The Davidson method has shown great success to approximate **exterior eigenpairs** of **diagonally dominant**, *but not diagonal*, matrices $A$.
  Indeed, **if $A$ is diagonal**, then

$$t = (D_A - \lambda I_n)^{-1} r = (A - \lambda I_n)^{-1}(A - \lambda I_n)y = y \in \text{range}(Q_k)$$

  so that **the search space range$(Q_k)$ cannot be expanded with** $t$.
  Then the method **stagnates** and becomes unable to achieve convergence.
- If $D_A \propto I_n$, then the Davidson method is equivalent to Lanczos or Arnoldi.

Davidson E. R. (1975). The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices. J. Comput. Phys., 17, 87–94.

# Davidson method, cont'd$_2$

- Although the Davidson method was originaly applied to real symmetric matrices, it seamlessly applies to non-symmetric/Hermitian matrices.

  For **non-symmetric/Hermitian** matrices, the algorithm is as follows:

**Algorithm 1** Davidson: $(A, q, k) \mapsto (\lambda, y)$

1: Allocate memory for $Q_k, W_k \in \mathbb{F}^{n\times k}$ and $B_k \in \mathbb{F}^{k\times k}$
2: $q_1 := q/\|q\|_2$
3: **for** $j = 1, \ldots, k$ **do**
4: $\quad w_j := Aq_j$
5: $\quad B_k[1:j, j] := Q_j^H w_j$, $B_k[j, 1:j-1] := q_j^H W_{j-1}$
6: $\quad$ Solve for an exterior eigenpair $(\lambda, \hat{y})$ of $B_j$ $\quad \triangleright\ B_j := B_k[1:j, 1:j]$
7: $\quad y := Q_j \hat{y}$
8: $\quad r := Ay - \lambda y$
9: $\quad$ Solve for $t$ such that $(D_A - \lambda I_n)t = r$
10: $\quad t := \Pi^{(j)} t$ $\quad \triangleright\ \Pi^{(j)}$ is a projector onto range$(Q_j)^\perp$
11: $\quad q_{j+1} := t/\|t\|_2$

Davidson E. R. (1975). The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices. J. Comput. Phys., 17, 87–94.

# Davidson method, cont'd$_3$

- For **Hermitian matrices**, the algorithm is as follows:

**Algorithm 2** Davidson: $(A, q, k) \mapsto (\lambda, y)$

1: Allocate memory for $Q_k \in \mathbb{F}^{n\times k}$ and $B_k \in \mathbb{F}^{k\times k}$
2: $q_1 := q/\|q\|_2$
3: **for** $j = 1, \ldots, k$ **do**
4: $\quad w := Aq_j$
5: $\quad B_k[1:j, j] := Q_j^H w$ ▷ $B_k[j, 1:j-1] := B_k[1:j-1, j]^T$
6: $\quad$ Solve for an exterior eigenpair $(\lambda, \hat{y})$ of $B_j$ ▷ $B_j := B_k[1:j, 1:j]$
7: $\quad y := Q_j \hat{y}$
8: $\quad r := Ay - \lambda y$
9: $\quad$ Solve for $t$ such that $(D_A - \lambda I_n)t = r$
10: $\quad t := \Pi^{(j)} t$ ▷ $\Pi^{(j)}$ is a projector onto range$(Q_j)^\perp$
11: $\quad q_{j+1} := t/\|t\|_2$

Although $\Pi^{(j)}$ is most commonly defined as a MGS procedure, CGS2 can also be used to mitigate potential stagnation, see van der Vorst (2002).

Davidson E. R. (1975). The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real symmetric matrices. J. Comput. Phys., 17, 87–94.
van der Vorst, H. A. (2002). Computational methods for large eigenvalue problems.

# Generalized Davidson (GD) method

# Generalized Davidson method

- Some modificiations of the Davidson method introduced by Morgan & Scott (1986) as well as Crouzeix et al. (1994) lead to variants collectively referrred to as the **generalized Davidson (GD)** method.

  Morgan & Scott (1986):

  - A **general preconditioner** of the form $M - \vartheta I_n$ is used instead of the original $D_A - \lambda I_n$, **without requirement of positive-definiteness**.
  - The aim is for $M - \vartheta I_n$ to **approximate** $A - \lambda I_n$ while allowing for **fast evaluation of** $r \mapsto (M - \vartheta I_n)^{-1} r$.
  - The **preconditioner should not be too good** to avoid stagnation, i.e.,

$$M - \vartheta I_n = A - \lambda I_n \implies t := (M - \vartheta I_n)^{-1} r = y \in \mathsf{range}(Q_k).$$

    This **contradicts the common notion of preconditioner**.
  - Numerical results reported with significantly **improved convergence** behaviors when **letting** $M$ **be the tridiagonal form of** $A$ and $\vartheta := \lambda$.
  - Upon setting $\vartheta := \sigma$, one can **drive global convergence** toward some $\sigma$.

Morgan, R. B., & Scott, D. S. (1986). Generalizations of Davidson's method for computing eigenvalues of sparse symmetric matrices. SIAM Journal on Scientific and Statistical Computing, 7(3), 817-825.

# Generalized Davidson method, cont'd$_1$

Morgan (1991):

- **Harmonic Ritz** pairs are used to yield even **faster convergence towards interior eigeinpairs**.

Crouzeix et al. (1994):

- **Several eigenpairs** are sought **at the same time** and **several vectors** are **incorporated into each search space expansion**, leading to a **block implementation**.
- A **maximum search space dimension** is introduced, triggering periodic **restarts** of the iterative method.
- **Restarting** of the GD method **renders convergence dependent on positive-definiteness of the preconditioners**.

Sadkane (1993):

- **Extension** of the GD method **to real non-symmetric matrices**.

Morgan, R. B. (1991). Computing interior eigenvalues of large matrices. Linear Algebra and its Applications, 154, 289-309.

Crouzeix, M., Philippe, B., & Sadkane, M. (1994). The Davidson method. SIAM Journal on Scientific Computing, 15(1), 62-76.

Sadkane, M. (1993). Block-Arnoldi and Davidson methods for unsymmetric large eigenvalue problems. Numerische Mathematik, 64, 195-211.

# Generalized Davidson method, cont'd$_2$

- For a **harmonic Ritz** pair $(\lambda, y)$ with respect to a search space range$(Q_k)$, there exists $\hat{y} \in \mathbb{F}^k$ such that $y = Q_k \hat{y}$ and

$$G_1 \hat{y} = \lambda G_2^H \hat{y}$$

where $G_1 := ((A - \sigma I_n)Q_k)^H (A - \sigma I_n)Q_k$ and $G_2 := Q_k^H (A - \sigma I_n)Q_k$. If $A$ is non-Hermitian, the harmonic GD method is as follows:

**Algorithm 3** Harmonic GD: $(A, q, \sigma, k) \mapsto (\lambda, y)$

1: Allocate memory for $Q_k, W_k \in \mathbb{F}^{n \times k}$ and $G_1, G_2 \in \mathbb{F}^{k \times k}$
2: $q_1 := q/\|q\|_2$
3: **for** $j = 1, \ldots, k$ **do**
4: $\quad w_j := (A - \sigma I_n) q_j$
5: $\quad G_1[1:j, j] := W_j^H w_j$ ▷ $G_1[j, 1:j-1] := G_1[1:j-1, j]^H$
6: $\quad G_2[1:j, j] := Q_j^H w_j$, $G_2 := [j, 1:j-1] := q_j^H W_{j-1}$
7: $\quad$ Solve for eigenpair $(\lambda, \hat{y})$ of $(G_1[1:j, 1:j], G_2[1:j, 1:j]^H)$ closest to $\sigma$
8: $\quad y := Q_j \hat{y}$, $\rho := \sigma + \hat{y}^H G_2 \hat{y}$, $r := Ay - \rho y$
9: $\quad$ Solve for $t$ such that $(M - \rho I_n)t = r$
10: $\quad t := \Pi^{(j)} t$ ▷ $\Pi^{(j)}$ is a projector onto range$(Q_j)^\perp$
11: $\quad q_{j+1} := t/\|t\|_2$

# Jacobi methods

# Orthogonal complement corrections

- **Jacobi orthogonal complement correction (JOCC)**:
  Jacobi (1845) considered an eigenvalue problem as a linear system of equations for which an iterative solver, e.g., Jacobi iteration, is used as a means to generate a **sequence of orthogonal complement corrections to a given approximate eigenvector**.
  - Suppose we have a **strongly diagonally dominant** matrix $A$, of which $\alpha := a_{11}$ is the largest element.
  - Then $(\alpha, e_1)$ is an **approximation** of the **largest eigenpair** $(\theta, z)$ of $A$.
  - In matrix notation, the JOCC approach is as follows. Consider

$$A\left(e_1 + \begin{bmatrix} 0 \\ w \end{bmatrix}\right) = \theta\left(e_1 + \begin{bmatrix} 0 \\ w \end{bmatrix}\right)$$

$$\begin{bmatrix} \alpha & c^T \\ b & F \end{bmatrix}\left(e_1 + \begin{bmatrix} 0 \\ w \end{bmatrix}\right) = \theta\left(e_1 + \begin{bmatrix} 0 \\ w \end{bmatrix}\right)$$

  where $[0\, w^T]^T$ is an **orthogonal complement correction** to the approximate eigenvector $e_1$.

Jacobi, C.G.J. (1845), Ueber eine neue Auflosungsart der bei der Methode der kleinsten Quadrate vorkommende linearen Gleichungen, Astronom Nachr, 297-306.

# Orthogonal complement corrections, cont'd

- The eigenvalue problem leads to the following equations:

$$\lambda = \alpha + c^T w$$
$$(F - \lambda I_n)w = -b$$

  which Jacobi proceeded to solve with the following iteration:

$$\lambda_k := \alpha + c^T w_k$$
$$(D_F - \lambda_k I_n)w_{k+1} = (D_F - F)w_k - b$$

  with $w_1 := 0$.
- This later became known as the **Jacobi iteration**.
- Although $\lambda_k$ is not a Rayleigh-Ritz value, it is nevertheless an approximation of the largest eigenvalue $\theta$.

- As the **JOCC** approach is **best-suited for strongly diagonally dominant matrices**, Jacobi (1846) coupled this approach with a **set of rotations** whose application **makes the matrix more diagonally dominant**.

Jacobi, C.G.J. (1845), Ueber eine neue Auflosungsart der bei der Methode der kleinsten Quadrate vorkommende linearen Gleichungen, Astronom Nachr, 297-306.

Jacobi, C. G. J. (1846). Über ein leichtes Verfahren die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen. J. Reine Angew. Math., 30, 51–94.

# Jacobi-Davidson method

# JOCC by Sleijpen and van der Vorst (1996)

- Sleijpen and van der Vorst (1996) revisited the JOCC approach in the more general setting where an arbitrary iterate $y_j \in \mathbb{F}^n$ is known as an approximate eigenvector of $A \in \mathbb{F}^{n \times n}$.

  JOCC's adaptation to this more general context lies in setting a correction $\delta \in \mathbb{F}^n$ to $y_j$ such that

$$A(y_j + \delta) = \theta(y_j + \delta), \quad \delta \perp y_j \tag{1}$$

  where $\theta$ is the wanted eigenvalue of $A$.

  Eq. (1) can be decomposed into two parts, along and orthogonal to $y_j$:

  - First, the **part parallel to** $y_j$, given by

$$y_j y_j^H A(y_j + \delta) = \theta y_j y_j^H (y_j + \delta)$$

    is such that $\boxed{\vartheta_j + y_j^H A\delta = \theta}$ where $\boxed{\vartheta_j := y_j^H A y_j}$ is the **corrected eigenvalue estimate**.

Sleijpen, G. L., & Van der Vorst, H. A. (1996). A Jacobi–Davidson iteration method for linear eigenvalue problems. SIAM review, 42(2), 267-293.

# JOCC by Sleijpen and van der Vorst (1996), cont'd

- Second, the **part orthogonal to** $y_j$, given by

$$(I_n - y_j y_j^H)A(y_j + \delta) = \theta(I_n - y_j y_j^H)(y_j + \delta)$$

is such that

$$\begin{aligned}(I_n - y_j y_j^H)(A - \theta I_n)\delta &= (I_n - y_j y_j^H)(-Ay_j + \theta y_j) \\ &= -(I_n - y_j y_j^H)Ay_j \\ &= -(Ay_j - \vartheta_j y_j) =: -r_j.\end{aligned}$$

Since $\delta \perp y_j$, we have $\delta = (I_n - y_j y_j^H)\delta$, and we obtain:

$$(I_n - y_j y_j^H)(A - \theta I_n)(I_n - y_j y_j^H)\delta = -r_j$$

where $\theta$, which **is unknown**, is replaced by $\vartheta_j$ to yield the **Jacobi-Davidson correction equation** given by:

$$\boxed{(I_n - y_j y_j^H)(A - \vartheta_j I_n)(I_n - y_j y_j^H)\delta = -r_j}.$$

Sleijpen, G. L., & Van der Vorst, H. A. (1996). A Jacobi–Davidson iteration method for linear eigenvalue problems. SIAM review, 42(2), 267-293.

# Homework problems

# Homework problem

Turn in **your own** solution to **Pb. 25**:

**Pb. 25** Show that if a matrix $A \in \mathbb{F}^{n \times n}$ has constant diagonal components, i.e., $D_A \propto I_n$, then, assuming exact arithmetic, the Davidson method is equivalent to applying a Rayleigh-Ritz projection to a Krylov subspace generated by a Arnoldi procedure with the same starting vector $q$.

**Pb. 26** ***.