Orthogonalization

Consider
$\mathbf{V}_k = [\mathbf{v}_1 \,|\, \ldots \,|\, \mathbf{v}_k]$ and $\mathbf{W}_k = [\mathbf{w}_1 \,|\, \ldots \,|\, \mathbf{w}_k]$,
matrices of full rang (of rank k).

An **expansion vector** $\tilde{\mathbf{v}}$ for $\mathbf{V}_k$
is **orthogonalized w.r.t.** $\mathbf{W}_k$ by

$$\tilde{\mathbf{u}} = \tilde{\mathbf{v}} - \mathbf{V}_k\,(\mathbf{W}_k^*\mathbf{V}_k)^{-1}\,(\mathbf{W}_k^*\tilde{\mathbf{v}}),$$

select a $\gamma_k \neq 0$, $\quad \mathbf{v}_{k+1} = \tilde{\mathbf{u}}/\gamma_k$.

Then $\quad \mathbf{v}_{k+1} \perp \mathbf{w}_j \quad \forall j \leq k$
or, equivalently, $\quad \mathbf{W}_k^*\mathbf{v}_{k+1} = 0$.

Moreover, $\tilde{\mathbf{v}} = [\mathbf{V}_k \,|\, \mathbf{v}_{k+1}] \begin{bmatrix} \vec{h}_k \\ \gamma_k \end{bmatrix}$
with $\vec{h}_k := (\mathbf{W}_k^*\mathbf{V}_k)^{-1}\,(\mathbf{W}_k^*\tilde{\mathbf{v}})$

$\mathbf{V}_k$ & $\mathbf{W}_k$ matrices of full rank

$(\mathbf{V}_k, \mathbf{W}_k)$ are **bi-orthogonal** iff
$\mathbf{W}_k^*\mathbf{V}_k$ is a non-singular diagonal

$\mathbf{V}_k$ is **orthogonal** iff
$\mathbf{V}_k^*\mathbf{V}_k$ is a non-singular diagonal

$\mathbf{V}_k$ is ortho**normal** iff $\mathbf{V}_k^*\mathbf{V}_k = I$

How to evaluate $\quad \tilde{\mathbf{v}} - \mathbf{V}_k\,(\mathbf{W}_k^*\mathbf{V}_k)^{-1}\,(\mathbf{W}_k^*\tilde{\mathbf{v}})$ *?*

<u>Important case: $\mathbf{V}_k$ orthogonal</u>
How to evaluate $\quad \tilde{\mathbf{v}} - \mathbf{V}_k\, D_k^{-1}(\mathbf{V}_k^*\tilde{\mathbf{v}})$ *?*
where $\quad D_k := \mathbf{V}_k^*\mathbf{V}_k \quad$ is diagonal

$D_k = \mathbf{V}_k^*\mathbf{V}_k = \text{diag}(\sigma_1, \ldots, \sigma_k)$
How to evaluate $\quad \tilde{\mathbf{v}} - \mathbf{V}_k\, D_k^{-1}(\mathbf{V}_k^*\tilde{\mathbf{v}})$ *?*

Gram-Schmidt orthogonalization

```
for i = 1,...,k do
    τ_i = (ṽ, v_i),  β_i = τ_i/σ_i
end for
for i = 1,...,k do
    ṽ ← ṽ − β_i v_i
end for
```

$\vec{\tau} = \mathbf{V}_k^*\tilde{\mathbf{v}}$
$\vec{\beta} = D_k^{-1}\vec{\tau}$
$\tilde{\mathbf{v}} \leftarrow \tilde{\mathbf{v}} - \mathbf{V}_k\vec{\beta}$

<u>**Advantage**</u>

- Additional parallelism

<u>**Drawback**</u>

- May be unstable

Gram–Schmidt is unstable (if the vectors to be orthogonalized make small angles)

[Walker '92]

Classical Gram-Schmidt

(Classical) Gram–Schmidt is unstable
(if the vectors to be orthogonalized make small angles)

$\vec{\tau} = \mathbf{V}_k^*\tilde{\mathbf{v}}$
$\vec{\beta} = D_k^{-1}\vec{\tau}$
$\tilde{\mathbf{v}} \leftarrow \tilde{\mathbf{v}} - \mathbf{V}_k\vec{\beta}$

Compute all components, compute the projection on the span of all $\mathbf{v}_j$ *(sum of components) and subtract the projection.*

All errors are summed.

Modified Gram–Schmidt

```
for i = 1,...,k do
    τ_i = (ṽ, v_i),   β_i = τ_i/σ_i
    ṽ ← ṽ − β_i v_i
end for
```

Project and subtract the components one by one: compute one component and subtract that component, ...

Any error is projected and components are subtracted also.

Modified Gram-Schmidt is a less unstable than classical Gram-Schmidt. Its stability is satisfactory in practice for linear solvers.
As expensive as classical Gram-Schmidt (on non-parallel machines)

Repeated Gram-Schmidt

$\rho = (\tilde{\mathbf{v}}, \tilde{\mathbf{v}}),\quad \vec{\beta} = D_k^{-1}(\mathbf{V}_k^*\tilde{\mathbf{v}})$
$\tilde{\mathbf{v}} \leftarrow \tilde{\mathbf{v}} - \mathbf{V}_k\vec{\beta}$
$\sigma_{k+1} = (\tilde{\mathbf{v}}, \tilde{\mathbf{v}})$,
while $\sigma_k < \kappa\,\rho$ $\qquad$ e.g., $\kappa = 0.5$
$\quad \vec{\gamma} = D_k^{-1}(\mathbf{V}_k^*\tilde{\mathbf{v}})$
$\quad \tilde{\mathbf{v}} \leftarrow \tilde{\mathbf{v}} - \mathbf{V}_k\vec{\gamma}$
$\quad \vec{\beta} \leftarrow \vec{\beta} + \vec{\gamma}$
$\quad \rho \leftarrow \sigma_{k+1},\quad \sigma_{k+1} = (\tilde{\mathbf{v}}, \tilde{\mathbf{v}})$
end while

<u>In practice</u>:

- Appears to be stable
- At most 1 refinement step needed (Only a little more expensive than Gram-Schmidt)

<u>Theory</u>:
Proof for "once is enough" only if $k = 1$.

Householder reflections

+ Are very stable
− Twice as expensive as
classical or modified Gram-Schmidt.

[Gronopolous '92]

s-step approach

Work with blocks of s vectors:
Orthogonalize $\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_s$ w.r.t. $\mathbf{V}_k = [\mathbf{v}_1 \,|\, \ldots \,|\, \mathbf{v}_k]$.

+ More parallelism can be exploited
− Can be very unstable

Krylov basis

$\mathbf{v}_1, \ldots, \mathbf{v}_k$ is a **Krylov basis** of $\mathcal{K}_k(\mathbf{A}; \mathbf{r}_0)$ if
$\mathcal{K}_j(\mathbf{A}; \mathbf{r}_0) = \text{span}\big(\mathbf{v}_1, \ldots, \mathbf{v}_j\big)$ for all $j \leq k$

Note that then $\quad \mathcal{K}_k(\mathbf{A}; \mathbf{r}_0) = \mathcal{K}_k(\mathbf{A}; \mathbf{v}_1)$.
Therefore, when referring to $\mathbf{A}$*, we will not specify the Krylov subspace explicitly.*

<u>**Theorem**</u>. If $\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}$ is a Krylov basis
then $\quad \boxed{\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k}$

where $\mathbf{V}_j := [\mathbf{v}_1 \,|\, \ldots \,|\, \mathbf{v}_j]$ and
$\underline{H}_k$ is a $(k+1) \times k$ (upper) Hessenberg.
$\underline{H}_k$ is **irreducible**, i.e. $h_{j+1\,j} \neq 0 \quad \forall j$.

H_k is the $k \times k$ upper block of $\underline{H}_k$.

Arnoldi

Arnoldi expands recursively an orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of $\mathcal{K}_k(\mathbf{A}; \mathbf{r}_0)$ ($\rho \mathbf{v}_1 = \mathbf{r}_0$) by orthogonalizing $\mathbf{A}\mathbf{v}_k$ w.r.t. $\mathbf{v}_1, \ldots, \mathbf{v}_k$:

$$\tilde{\mathbf{v}} = \mathbf{A}\mathbf{v}_k - \sum_{j=1}^{k} h_{jk}\, \mathbf{v}_j \quad \text{with} \quad h_{jk} = (\mathbf{A}\mathbf{v}_k, \mathbf{v}_j)$$

$$\mathbf{v}_{k+1} = \tilde{\mathbf{v}}/h_{k+1\,k} \quad \text{with} \quad h_{k+1\,k} = \|\tilde{\mathbf{v}}\|$$

<u>Theorem</u>

Then $\boxed{\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\, \underline{H}_k}$ where

$$\mathbf{V}_k := [\mathbf{v}_1 \,|\, \ldots \,|\, \mathbf{v}_k] \quad \text{and} \quad \underline{H}_k := \begin{bmatrix} h_{11} & \ldots & h_{1k} \\ h_{21} & & \\ & \ddots & \vdots \\ 0 & & h_{k+1\,k} \end{bmatrix}$$

$\mathbf{V}_k$ is orthonormal: $\mathbf{V}_k^* \mathbf{V}_k = I$,

$\underline{H}_k$ is $(k+1) \times k$ irreducible upper Hessenberg

[Arnoldi '51]

Arnoldi's expansion

Choose $\mathbf{r}_0$
$k = 1$
$\rho = \|\mathbf{r}_0\|$, $\mathbf{v}_1 = \mathbf{r}_0/\rho$
$\mathbf{V}_1 = [\mathbf{v}_1]$, $\underline{H}_0 = [\;]$
while ... do
  $\tilde{\mathbf{v}} \leftarrow \mathbf{A}\, \mathbf{v}_k$
  $\vec{h}_k = \mathbf{V}_k^* \tilde{\mathbf{v}}$, $\tilde{\mathbf{v}} \leftarrow \tilde{\mathbf{v}} - \mathbf{V}_k \vec{h}_k$
  $h_{k+1\,k} = \|\tilde{\mathbf{v}}\|$, $\mathbf{v}_{k+1} = \tilde{\mathbf{v}}/h_{k+1\,k}$
  $\mathbf{V}_{k+1} = [\mathbf{V}_k \,|\, \mathbf{v}_{k+1}]$, $\underline{H}_k = \begin{bmatrix} \underline{H}_{k-1} & \vec{h}_k \\ 0 \ldots 0 & h_{k+1\,k} \end{bmatrix}$
  $k \leftarrow k + 1$
end while

Arnoldi with Householder r.

Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k$ orthonormal,
$\mathbb{H}_2, \ldots, \mathbb{H}_k$ Householder reflections s.t.

$$\mathbb{H}_k \cdots \mathbb{H}_3 \cdot \mathbb{H}_2 \cdot [\,\mathbf{A}\mathbf{v}_1 \,|\, \ldots \,|\, \mathbf{A}\mathbf{v}_k\,] = \begin{bmatrix} R_k \\ \mathbf{0} \end{bmatrix}$$

with R_k upper triangular

$\tilde{\mathbf{v}} \leftarrow \mathbb{H}_k \cdots \mathbb{H}_2 \cdot (\mathbf{A}\mathbf{v}_k)$
$\mathbb{H}_{k+1}$ reflects $(\tilde{\mathbf{v}})_{j>k}$ on $\text{span}\big(e_{k+1}\big)$
$\mathbf{v}_{k+1} = \mathbb{H}_2 \cdots \mathbb{H}_{k+1}\, e_{k+1}$
$R_{k+1} = \left[\begin{bmatrix} R_k \\ 0 \ldots 0 \end{bmatrix} (\mathbb{H}_{k+1}\, \tilde{\mathbf{v}})_{j \le k+1} \right]$

Then $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\, \underline{H}_k$ with $\mathbf{V}_j = [\mathbf{v}_1 \,|\, \ldots \,|\, \mathbf{v}_j]$
and $\underline{H}_k\, e_j = R_{k+1} e_{j+1} \quad (j = 1, \ldots, k)$

$\mathbf{y} = (y_1, \ldots, y_n)^T$. With $\mathbf{w} := \|\mathbf{y}\|\, e_k + \text{sign}(y_k)\, \mathbf{y}$,
$\mathbb{H} = \mathbf{I} - 2\frac{\mathbf{w}\,\mathbf{w}^*}{\mathbf{w}^*\mathbf{w}}$, reflects $\mathbf{y}$ to $\text{span}(e_k)$.
$(\mathbf{y})_{j \le k} := (y_1, \ldots, y_k)^T$ & $(\mathbf{y})_{j>k} := (0, \ldots, 0, y_{k+1}, \ldots, y_n)^T$

s-step Arnoldi

Expand an orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ (with $k = s\ell$) of $\mathcal{K}_k(\mathbf{A}; \mathbf{r}_0)$ by

$$\tilde{\mathbf{v}}_1 = \mathbf{A}\, \mathbf{v}_k, \quad \tilde{\mathbf{v}}_2 = \mathbf{A}\tilde{\mathbf{v}}_1, \; \ldots, \quad \tilde{\mathbf{v}}_s = \mathbf{A}\tilde{\mathbf{v}}_{s-1}$$

Orthogonalize these s-vector w.r.t. $\mathbf{v}_1, \ldots, \mathbf{v}_k$

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\, \underline{H}_k$$

is an **Arnoldi expansion of order** k,
i.e., $\mathbf{V}_{k+1}^* \mathbf{V}_{k+1} = I$, $\underline{H}_k$ irreducible Hessenberg

---

Lanczos

H_k the upper $k \times k$ block of $\underline{H}_k$.

<u>Theorem</u>
H_k is symmetric tridiagonal if $\mathbf{A}^* = \mathbf{A}$

Lanczos exploits this fact: it expands recurs. an orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of $\mathcal{K}_k(\mathbf{A}, \mathbf{r}_0)$ by:

$$\tilde{\mathbf{v}} = \mathbf{A}\, \mathbf{v}_k - \alpha_k\, \mathbf{v}_k - \beta_{k-1}\, \mathbf{v}_{k-1}$$
$$\text{with } \alpha_k = (\mathbf{A}\mathbf{v}_k, \mathbf{v}_k)$$
$$\mathbf{v}_{k+1} = \tilde{\mathbf{v}}/\beta_k \quad \text{with} \quad \beta_k = \|\tilde{\mathbf{v}}\|.$$

Then $\boxed{\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\, \underline{T}_k}$ where

$\underline{T}_k$ is $(k+1) \times k$ irreducible tridiagonal

[Lanczos '50]

Lanczos

Choose $\mathbf{r}_0$
$k = 1$
$\rho = \|\mathbf{r}_0\|$, $\mathbf{v}_1 = \mathbf{r}_0/\rho$
$\beta_0 = 0$, $\mathbf{v}_0 = \mathbf{0}$
while ... do
  $\tilde{\mathbf{v}} \leftarrow \mathbf{A}\, \mathbf{v}_k - \beta_{k-1}\, \mathbf{v}_{k-1}$
  $\alpha_k = \mathbf{v}_k^* \tilde{\mathbf{v}}$, $\tilde{\mathbf{v}} \leftarrow \tilde{\mathbf{v}} - \alpha_k\, \mathbf{v}_k$
  $\beta_k = \|\tilde{\mathbf{v}}\|$, $\mathbf{v}_{k+1} = \tilde{\mathbf{v}}/\beta_k$
  $k \leftarrow k + 1$
end while

Lanczos versus Arnoldi: $\mathbf{A}^* = \mathbf{A}$

Arnoldi: stable (very stable with repeated GS)

Lanczos: very unstable
but, very surprisingly, still usable

Arnoldi: costs (storage & computation)
grow with k
Lanczos: costs constant and small per step

<u>Alternative approaches</u>

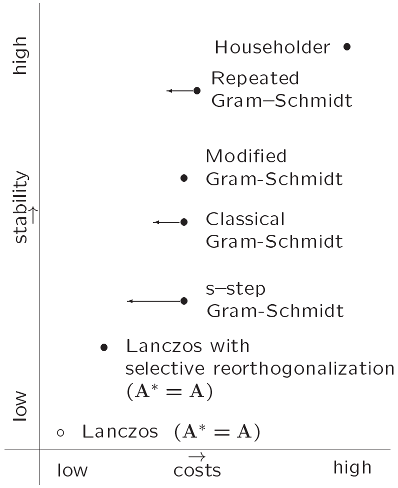
if Lanczos needs more stability

- **Selective Reorthogonalization**: orthogonalize also with respect to an appropriate part of the "old" space
- **Complete reorthogonalization**: Arnoldi but retain only tridiagonal part of $\underline{H}_k$.

The error in Lanczos process propagates in specific directions

plot $\text{fl}(\mathbf{V}_k)^* \text{fl}(\mathbf{V}_k)$ and $\mathbb{V}^*(\text{fl}(\mathbf{V}_k) - \mathbf{V}_k)$

Properties orthogonalization



• proportional with k, ∘ fixed per step

⟵ Improvements on parallel machines

The above figure shows the $\|I - \mathbf{V}_k^*\mathbf{V}_k\|$ (on ^{10}log-scale) for several orthogonalization techniques

Minimal residual methods

All minimal residual methods require the computation of an orthogonal basis.

for $\mathcal{K}_k(\mathbb{A};\mathbf{r}_0)$, $\mathbf{v}_1 = \mathbf{r}_0$ (GMRES, FOM, *Arnoldi*)

for $\mathbb{A}\mathcal{K}_k(\mathbb{A};\mathbf{r}_0)$, $\mathbf{v}_1 = \mathbb{A}\mathbf{r}_0$ (ORTHOMIN, GCR, ORTHODIR)

ORTHOMIN, GCR expand with $\mathbb{A}\,\mathbf{r}_k$ (or $\mathbb{A}\mathbf{K}_k^{-1}\mathbf{r}_k$)
- − *may break down,*
- + *allow varying preconditioning*

GMRES, ORTHODIR expand with $\mathbb{A}\,\mathbf{c}_{k-1}$ $(\mathbb{A}\mathbf{v}_k)$
- + *robust,*
- − *allows only fixed preconditioning*

GMRES, FOM, *Arnoldi* normalize the orthogonalizes expansion vector

How to evaluate $\tilde{\mathbf{v}} - \mathbf{V}_k\,(\mathbf{W}_k^*\mathbf{V}_k)^{-1}(\mathbf{V}_k^*\tilde{\mathbf{v}})$*?*

Theorem.
Modified Gram-Schmidt is possible iff $M_k := \mathbf{W}_k^*\mathbf{V}_k$ is upper triangular.

[Lanczos '50]

Bi–Lanczos

If $\mathbf{v}_1,\ldots,\mathbf{v}_k$ is a Krylov basis of $\mathcal{K}_k(\mathbf{A};\mathbf{r}_0)$
and $\mathbf{w}_1,\ldots,\mathbf{w}_k$ is a Krylov bas. of $\mathcal{K}_k(\mathbf{A}^*;\tilde{\mathbf{r}}_0)$,
then **Bi-Lanczos** expands recursively
- $\mathbf{v}_1,\ldots,\mathbf{v}_k$ by $\mathbf{A}\mathbf{v}_k$ orthogonalized w.r.t. $\mathbf{w}_1,\ldots,\mathbf{w}_k$,
- and $\mathbf{w}_1,\ldots,\mathbf{w}_k$ by $\mathbf{A}^*\mathbf{w}_k$ orthogonalized w.r.t. $\mathbf{v}_1,\ldots,\mathbf{v}_k$.

Put $\mathbf{V}_j = [\mathbf{v}_1\,|\ldots|\,\mathbf{v}_j]$ and $\mathbf{W}_j = [\mathbf{w}_1\,|\ldots|\,\mathbf{w}_j]$.
The bi-Lanczos process does not **break down** if $D_{k+1} := \mathbf{W}_{k+1}^*\mathbf{V}_{k+1}$ is non-singular diagonal.
$\underline{H}_k$ and $\underline{F}_k$ are Hessenberg matrices such that

$$\mathbf{A}\,\mathbf{V}_k = \mathbf{V}_{k+1}\,\underline{H}_k \;\&\; \mathbf{A}^*\,\mathbf{W}_k = \mathbf{W}_{k+1}\,\underline{F}_k$$

Theorem. $H_k = D_k^{-1}\,F_k^*\,D_k$ is tridiagonal, if the bi-Lanczos process does not break down.

Proof. $F_k^* D_k = \mathbf{W}_k^*\,\mathbf{A}\,\mathbf{V}_k = D_k H_k$

Bi-Lanczos

Select a $\mathbf{r}_0$, and a $\tilde{\mathbf{r}}_0$
$k = 1$
$\mathbf{v}_1 = \mathbf{r}_0/\|\mathbf{r}_0\|$, $\mathbf{v}_0 = \mathbf{0}$, $\mathbf{w}_1 = \tilde{\mathbf{r}}_0/\|\tilde{\mathbf{r}}_0\|$, $\mathbf{w}_0 = \mathbf{0}$
$\gamma_0 = 0$, $\delta_0 = 1$, $\tilde{\gamma}_0 = 0$, $\tilde{\delta}_0 = 1$
while ... do
  $\delta_k = \mathbf{w}_k^*\mathbf{v}_k$,
  $\tilde{\mathbf{v}} = \mathbf{A}\,\mathbf{v}_k$, $\tilde{\mathbf{w}} = \mathbf{A}^*\,\mathbf{w}_k$
  $\beta_k = \overline{\tilde{\gamma}}_{k-1}\delta_k/\delta_{k-1}$, $\tilde{\beta}_k = \overline{\gamma}_{k-1}\overline{\delta}_k/\overline{\delta}_{k-1}$
  $\tilde{\mathbf{v}} \leftarrow \tilde{\mathbf{v}} - \beta_k\mathbf{v}_{k-1}$, $\tilde{\mathbf{w}} \leftarrow \tilde{\mathbf{w}} - \tilde{\beta}_k\mathbf{w}_{k-1}$
  $\alpha_k = \mathbf{w}_k^*\tilde{\mathbf{v}}/\delta_k$, $\tilde{\alpha}_k = \overline{\alpha}_k$
  $\tilde{\mathbf{v}} \leftarrow \tilde{\mathbf{v}} - \alpha_k\,\mathbf{v}_k$, $\tilde{\mathbf{w}} \leftarrow \tilde{\mathbf{w}} - \tilde{\alpha}_k\,\mathbf{w}_k$
  Select a $\gamma_k \neq 0$ and a $\tilde{\gamma}_k \neq 0$
  $\mathbf{v}_{k+1} = \tilde{\mathbf{v}}/\gamma_k$, $\mathbf{w}_{k+1} = \tilde{\mathbf{w}}/\tilde{\gamma}_k$,
  $k \leftarrow k+1$
  $\mathbf{V}_k = [\mathbf{V}_{k-1}\,|\,\mathbf{v}_k]$, $\mathbf{W}_k = [\mathbf{W}_{k-1}\,|\,\mathbf{w}_k]$
end while

Minimal residual methods

Optimality:

$\mathbf{y}_k \in \mathcal{K}_k(\mathbb{A};\mathbf{r}_0)$ solves

$$\min\{\|\mathbf{r}_0 - \mathbb{A}\mathbf{y}\| \mid \mathbf{y} \in \mathcal{K}_k(\mathbb{A};\mathbf{r}_0)\}$$

or, equivalently

$\mathbf{s}_k \in \mathbb{A}\mathcal{K}_k(\mathbb{A};\mathbf{r}_0)$ solves

$$\min\{\|\mathbf{r}_0 - \mathbf{s}\| \mid \mathbf{s} \in \mathbb{A}\mathcal{K}_k(\mathbb{A};\mathbf{r}_0)\}$$

• = Minimal residual methods

• ORTHOMIN *ORTHOgonal MINimization*

⇓ *stability*

Generalized Conjugate Residuals

• GCR ⇒ • efficient GCR ↔ • GMRES

Breakdown ⇒ • ORTHODIR *ORTHOgonal DIRections*

⇓ *efficiency*

$\boxed{\text{Truncated GCR}}$ & $\boxed{\text{Restarted } \textit{GCR}\text{, GMRES}}$

⇓ *fast convergence*

$\boxed{\text{GMRESR}}$ *GMRES Repeated*

⇓ *fast convergence*

$\boxed{\text{GCRO}}$ *GCR Orthogonalized*

• $\boxed{\text{GMRES}}$ Generalized Minimal RESiduals ↔ ∘ FOM Ful Orthogonalization Method

Minimal residual methods

Advantages:

- Krylov subspace methods.
- Optimality.
- 1 MV needed to increase the Krylov subspace by one dimension
- No knowledge (in advance) on (spectral) properties of $\mathbb{A}$ is needed.
- ∘ Approximations of certain eigenvalues can cheaply be computed from β_{ij}, σ_j and α_j.

Drawbacks:

- Growing computational costs if the number of iteration steps increases, # DOTs & # AXPYs proportional to the step number.
- Growing memory requirement if the number of iteration steps increases, # large vectors is proportional to the step number.
- Any step requires DOTs.

$$\|\mathbf{r}_k\| = \min\{\|\mathbf{r}_0 - \mathbb{A}\mathbf{y}\| \mid \mathbf{y} \in \mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)\}$$

GMRES

If $\mathbb{A}\mathbf{V}_k = \mathbf{V}_{k+1}\underline{H}_k$ is the Arnoldi expansion of order k with $\mathbf{r}_0 = \rho\,\mathbf{v}_1$ then

$$\begin{aligned}\|\mathbf{r}_k\| &= \min_{\vec{\gamma}} \|\rho\,\mathbf{v}_1 - \mathbb{A}\,\mathbf{V}_k\,\vec{\gamma}\| \\ &= \min_{\vec{\gamma}} \|\rho\,\mathbf{v}_1 - \mathbf{V}_{k+1}\underline{H}_k\,\vec{\gamma}\| \\ &= \min_{\vec{\gamma}} \|\rho\,e_1 - \underline{H}_k\,\vec{\gamma}\| \quad (\star)\end{aligned}$$

If $\vec{\gamma}_k$ solves $(\star)$ then

$$\mathbf{x}_k = \mathbf{x}_0 + \mathbf{K}^{-1}\,\mathbf{V}_k\,\vec{\gamma}_k$$

Per step (neglecting operations of dimension k)

work: 1 MV, $k+1$ DOT, $k+1$ AXPY

memory: $\mathbf{x}_0$, $(k+1)\times\ \mathbf{v}_j$

Only once: $k+1$ AXPY.

[Saad Schultz '86]

GMRES

Choose $\mathbf{x}_0$

$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$

$\rho_0 = \rho = \|\mathbf{r}_0\|$, $\mathbf{v}_1 = \mathbf{r}_0/\rho$

$k = 0$, $\mathbf{V}_1 = [\mathbf{v}_1]$, $\underline{H}_0 = [\ ]$

while $\rho_{k-1} > \mathit{tol}$ do

  $k \leftarrow k+1$

  $\tilde{\mathbf{v}} \leftarrow \mathbf{K}^{-1}\mathbf{v}_k$, $\tilde{\mathbf{v}} \leftarrow \mathbf{A}\,\tilde{\mathbf{v}}$

  $\vec{h}_k = \mathbf{V}_k^*\,\tilde{\mathbf{v}}$, $\tilde{\mathbf{v}} \leftarrow \tilde{\mathbf{v}} - \mathbf{V}_k\,\vec{h}_k$

  $h_{k+1\,k} = \|\tilde{\mathbf{v}}\|$, $\mathbf{v}_{k+1} = \tilde{\mathbf{v}}/h_{k+1\,k}$

  $\mathbf{V}_{k+1} = [\mathbf{V}_k|\mathbf{v}_{k+1}]$, $\underline{H}_k = \begin{bmatrix} \underline{H}_{k-1} & \vec{h}_k \\ 0\ldots 0 & h_{k+1\,k}\end{bmatrix}$

  Compute $\rho_k = \min\|\rho\,e_1 - \underline{H}_k\,\vec{\gamma}\|$

end while

Solve $\min_{\vec{\gamma}}\|\rho\,e_1 - \underline{H}_k\,\vec{\gamma}\|$ $(\star)$

$\mathbf{x}_k = \mathbf{x}_0 + \mathbf{K}^{-1}(\mathbf{V}_k\,\vec{\gamma}_k)$ with $\vec{\gamma}_k$ solution $(\star)$

Solve $\min_{\vec{\gamma}}\|\rho\,e_1 - \underline{H}_k\,\vec{\gamma}\|$ $(\star)$

Consider the solution $\vec{\gamma}_k$ of $(\star)$

If $\underline{H}_k = Q_{k+1}\,\underline{R}_k$ is the QR-decomposition, that is Q_{k+1} is $(k+1)\times(k+1)$ unitary, $\underline{R}_k$ is $(k+1)\times k$ upper triangular, then

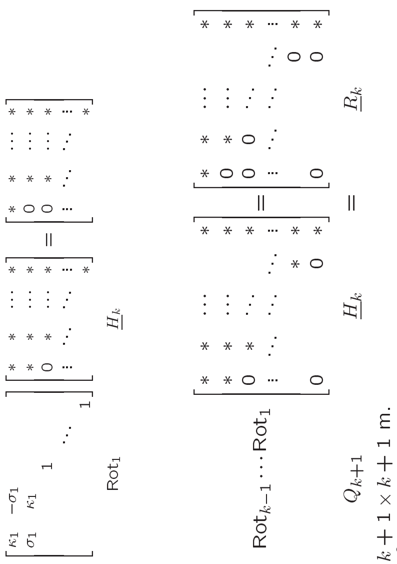
$$\|\rho\,e_1 - \underline{H}_k\,\vec{\gamma}_k\| = \|Q_{k+1}^*(\rho\,e_1) - \underline{R}_k\,\vec{\gamma}_k\|$$

Note that,

with $\vec{\zeta}_{k+1} := \begin{bmatrix}\tilde{\zeta}_k \\ \rho_k\end{bmatrix} := Q_{k+1}^*(\rho\,e_1)$

we have

$$\|\vec{\zeta}_{k+1} - \underline{R}_k\,\vec{\gamma}_k\| = |\rho_k| \quad\text{and}\quad \vec{\gamma}_k = R_k^{-1}\tilde{\zeta}_k$$



GMRES

While $\rho_k > \mathit{tol}$

  $k \leftarrow k+1$;

  Compute $\mathbf{v}_{k+1}$ and (h_{jk}) by one Arnoldi step;

  $\vec{\eta}_k := \text{Rot}_{k-1}\cdots\text{Rot}_1\begin{bmatrix} h_{1k} \\ \vdots \\ h_{k\,k}\end{bmatrix}$, $\underline{R}_k = \left[\begin{array}{c|c} R_{k-1} & \\ 0\cdots 0 & \vec{\eta}_k \\ \hline 0\cdots 0 & h_{k+1\,k}\end{array}\right]$

  Compute κ_k and σ_k; $\text{Rot}_k\,\underline{R}_k = \left[\begin{array}{c} R_k \\ \hline 0\cdots 0\end{array}\right]$;

  $\vec{\zeta}_{k+1} = \text{Rot}_k\begin{bmatrix}\vec{\zeta}_k \\ 0\end{bmatrix} = \begin{bmatrix}\tilde{\zeta}_k \\ \rho_k\end{bmatrix}$ (then $\rho_k = \rho_{k-1}\,\sigma_k$);

end while

Solve $R_k\,\vec{\gamma}_k = \tilde{\zeta}_k$ by backward substitution and compute $x_k = x_0 + \mathbf{V}_k\,\vec{\gamma}_k$

$$\min\{\|\mathbf{r}_0 - \mathbf{s}\| \mid \mathbf{s} \in \mathbb{A}\mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)\}$$

$\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{k-1}$ such that

$\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{k-1}$ is an orthogonal basis of $\mathbb{A}\mathcal{K}_k(\mathbb{A}; \mathbf{r}_0)$

where $\mathbf{c}_j = \mathbf{A}\mathbf{u}_j$ for $j = 0, 1, \ldots, k-1$

How to select the expansion vector for $\{\mathbf{c}_j\}$?

How to compute the $\mathbf{c}_k$?

Note that:

$$\mathbf{r}_k \in \mathcal{K}_k(\mathbb{A}; \mathbf{r}_0) \quad\Rightarrow\quad \mathbb{A}\mathbf{r}_k \in \mathbb{A}\mathcal{K}_k(\mathbb{A}; \mathbf{r}_0).$$

$$\mathbf{c}_k = \mathbb{A}\mathbf{r}_k - \sum_{j=0}^{k-1}\frac{(\mathbb{A}\mathbf{r}_k, \mathbf{c}_j)}{(\mathbf{c}_j, \mathbf{c}_j)}\mathbf{c}_j \quad\text{and}$$

$$\mathbf{u}_k = \mathbf{K}^{-1}\mathbf{r}_k - \sum_{j=0}^{k-1}\frac{(\mathbb{A}\mathbf{r}_k, \mathbf{c}_j)}{(\mathbf{c}_j, \mathbf{c}_j)}\mathbf{u}_j$$

ORTHOMIN uses Classical Gram-Schmidt

GCR uses modified Gram-Schmidt

[Vinsome '76]

ORTHOMIN

Choose $\mathbf{x}_0$

$k = 0$

$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$

while $\|\mathbf{r}_k\| > \mathit{tol}$ do

  solve $\mathbf{u}$ from $\mathbf{K}\,\mathbf{u} = \mathbf{r}_k$

  $\mathbf{c} \leftarrow \mathbf{A}\,\mathbf{u}$

  for $i = 0, \ldots, k-1$ do

    $\tau_{ik} = (\mathbf{c}, \mathbf{c}_i)$, $\beta_{ik} = \tau_{ik}/\sigma_i$

  end for

  for $i = 0, \ldots, k-1$ do

    $\mathbf{u} \leftarrow \mathbf{u} - \beta_{ik}\,\mathbf{u}_i$

    $\mathbf{c} \leftarrow \mathbf{c} - \beta_{ik}\,\mathbf{c}_i$

  end for

  $\mathbf{u}_k = \mathbf{u}$, $\mathbf{c}_k = \mathbf{c}$

  $\sigma_k = (\mathbf{c}_k, \mathbf{c}_k)$, $\rho_k = (\mathbf{r}_k, \mathbf{c}_k)$, $\alpha_k = \rho_k/\sigma_k$

  $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\,\mathbf{u}_k$

  $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\,\mathbf{c}_k$

  $k \leftarrow k+1$

end while

[Elman '82]

GCR

Choose $\mathbf{x}_0$
$k=0$
$\mathbf{r}_0=\mathbf{b}-\mathbf{A}\mathbf{x}_0$
while $\|\mathbf{r}_k\|>tol$ do
  solve $\mathbf{u}$ from $\mathbf{K}\,\mathbf{u}=\mathbf{r}_k$
  $\mathbf{c}\leftarrow\mathbf{A}\,\mathbf{u}$
  for $i=0,\ldots,k-1$ do
    $\tau_{ik}=(\mathbf{c},\mathbf{c}_i)$, $\beta_{ik}=\tau_{ik}/\sigma_i$
    $\mathbf{u}\leftarrow\mathbf{u}-\beta_{ik}\,\mathbf{u}_i$
    $\mathbf{c}\leftarrow\mathbf{c}-\beta_{ik}\,\mathbf{c}_i$
  end for
  $\mathbf{u}_k=\mathbf{u}$, $\mathbf{c}_k=\mathbf{c}$
  $\sigma_k=(\mathbf{c}_k,\mathbf{c}_k)$, $\rho_k=(\mathbf{r}_k,\mathbf{c}_k)$, $\alpha_k=\rho_k/\sigma_k$
  $\mathbf{x}_{k+1}=\mathbf{x}_k+\alpha_k\,\mathbf{u}_k$
  $\mathbf{r}_{k+1}=\mathbf{r}_k-\alpha_k\,\mathbf{c}_k$
  $k\leftarrow k+1$
end while

Savings in GCR

Proposition. $\mathbf{c}_k=\mathbf{A}\,\mathbf{u}_k$

∘ Useful for the k's for which
k AXPYs are more expensive than 1 MV.

- **Efficient GCR**
  Compute only the $\mathbf{c}_j$ and $\mathbf{r}_j$
  Store only β_{ij}, α_j, $\mathbf{c}_j$.
  If $\|\mathbf{r}_k\|$ is small enough,
  Compute $\mathbf{x}_k$ from
  the **implicit** reconstruction of
  the $\mathbf{r}_j$ and the $\mathbf{u}_j$

solve $\mathbf{u}$ from $\mathbf{K}\mathbf{u}=\mathbf{r}_k$

$$\mathbf{u}_k=\mathbf{u}-\sum_{i=0}^{k-1}\beta_{ik}\,\mathbf{u}_i$$

$$\mathbf{x}_{k+1}=\mathbf{x}_k+\alpha_k\,\mathbf{u}_k$$

Matrix representation

$\mathbf{K}^{-1}\mathbf{R}_{k+1}=\mathbf{U}_{k+1}\,B_{k+1}$, where

$\mathbf{R}_{k+1}:=[\,\mathbf{r}_0\,|\ldots|\,\mathbf{r}_k\,]$ & $\mathbf{U}_{k+1}:=[\,\mathbf{u}_0\,|\ldots|\,\mathbf{u}_k\,]$

$$B_{k+1}:=\begin{bmatrix}1&\beta_{01}&\beta_{02}&\ldots&\beta_{0k}\\0&1&\beta_{12}&\ldots&\beta_{1k}\\0&0&1&\ddots&\vdots\\\vdots&&\ddots&\ddots&\end{bmatrix}$$

$$\mathbf{x}_{k+1}=\mathbf{x}_k+\alpha_k\,\mathbf{u}_k\quad\Rightarrow\quad\mathbf{x}_{k+1}=\mathbf{x}_0+\mathbf{U}_{k+1}\begin{bmatrix}\alpha_0\\\vdots\\\alpha_k\end{bmatrix}$$

$\mathbf{x}_{k+1}=\mathbf{x}_0+\mathbf{K}^{-1}\mathbf{y}_k$ where $\mathbf{y}_k=\mathbf{R}_{k+1}(B_{k+1}^{-1}\,\vec{\alpha}_{k+1})$

[Sleijpen van der Vorst '93]

An efficient form of GCR

Choose $\mathbf{x}_0$
$k=0$, $\mathbf{r}_0=\mathbf{b}-\mathbf{A}\mathbf{x}_0$
while $\|\mathbf{r}_k\|>tol$ do
  solve $\mathbf{u}$ from $\mathbf{K}\,\mathbf{u}=\mathbf{r}_k$
  $\mathbf{c}\leftarrow\mathbf{A}\,\mathbf{u}$
  for $i=0,\ldots,k-1$ do
    $\tau_{ik}=(\mathbf{c},\mathbf{c}_i)$, $\beta_{ik}=\tau_{ik}/\sigma_i$
    $\mathbf{c}\leftarrow\mathbf{c}-\beta_{ik}\,\mathbf{c}_i$
  end for
  $\mathbf{c}_k=\mathbf{c}$,
  $\sigma_k=(\mathbf{c}_k,\mathbf{c}_k)$, $\rho_k=(\mathbf{r}_k,\mathbf{c}_k)$, $\alpha_k=\rho_k/\sigma_k$
  $\mathbf{r}_{k+1}=\mathbf{r}_k-\alpha_k\,\mathbf{c}_k$
  $k\leftarrow k+1$
end while
solve $\vec{\gamma}_k$ from $B_k\vec{\gamma}_k=\vec{\alpha}_k$
solve $\mathbf{y}$ from $\mathbf{K}\mathbf{y}=\mathbf{R}_k\vec{\gamma}_k$
$\mathbf{x}_k=\mathbf{x}_0+\mathbf{y}$

$$\mathbf{r}_{k+1}=\mathbf{r}_k-\alpha_k\,\mathbf{c}_k=\mathbf{r}_0-\sum_{i=0}^{k}\alpha_i\,\mathbf{c}_i=\mathbf{r}_0-\mathbf{C}_{k+1}\,\vec{\alpha}_{k+1}$$

Properties (efficient) GCR

Advantages:

- (One of) the most efficient minimal residual methods. 1 MV to increase the dimension of the Krylov subspace by 1.
In step k: $k+1$ DOT, k AXPY
$k+2$ large vectors to be stored

∘ Stability is satisfactory by the use of modified Gram-Schmidt to obtain orthogonal basis $\mathbf{c}_0,\ldots,\mathbf{c}_{k-1}$ of $\mathbb{A}\mathcal{K}_k(\mathbb{A};\mathbf{r}_0)$.

Drawbacks:

- Costs grow per iteration step.

∘ Not robust: iteration process may break down ($\rho_k=0$ $\Rightarrow$ $\mathbf{c}_{k+1}=0$)

∘ The approximation $\mathbf{x}_k$ is only known at the end.

Breakdown in GCR

Example: $\mathbf{A}=\mathbb{A}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$, $\mathbf{b}=\begin{bmatrix}1\\0\end{bmatrix}$

If $\mathbf{x}_0=0$ then $\mathbf{r}_0=\mathbf{b}=e_1$,
$\mathbf{c}_0=\mathbf{A}\mathbf{r}_0=e_2$ and $\rho_0=(\mathbf{r}_0,\mathbf{c}_0)=0$.

Remedy: Replace (in GCR)

solve $\mathbf{u}$ from $\mathbf{K}\mathbf{u}=\mathbf{r}_k$, $\mathbf{c}=\mathbf{A}\mathbf{u}$

by (ORTHODIR)

solve $\mathbf{u}$ from $\mathbf{K}\mathbf{u}=\mathbf{c}_{k-1}$, $\mathbf{c}=\mathbf{A}\mathbf{u}$

— If GCR does not break down then GCR $\mathbf{c}_j$ and ORTHODIR $\mathbf{c}_j$ are equal (except for a scalar multiple).

— ORTHODIR does not break down.

Explanation:

$\mathbf{r}_k=\mathbf{r}_0+\gamma_1\,\mathbb{A}\mathbf{r}_0+\ldots+\gamma_k\,\mathbb{A}^k\mathbf{r}_0$
γ_k may be 0; control on γ_0: $\gamma_0=1$

in ORTHODIR:
$\mathbf{c}_{k-1}=\gamma_1\,\mathbb{A}\mathbf{r}_0+\ldots+\gamma_{k-1}\,\mathbb{A}^{k-1}\mathbf{r}_0+\mathbb{A}^k\mathbf{r}_0$
control on γ_k: $\gamma_k=1$

As GCR, ORTHODIR can be reformulated in an efficient form (no breakdown & optimal efficient in the class of minimal residual methods)

GMRES & efficient ORTHODIR

GMRES has the same properties (**+ and −**) as efficient ORTHODIR

Orthogonal basis for:
— $\mathcal{K}_k(\mathbb{A};\mathbf{r}_0)$ in GMRES
— $\mathbb{A}\,\mathcal{K}_k(\mathbb{A};\mathbf{r}_0)$ in ORTHODIR (GCR)
produced by (class.) Gram-Schmidt:
$\mathbf{c}_0=\mathbf{r}_0$ (GMRES), $\mathbf{c}_0=\mathbb{A}\mathbf{r}_0$ (ORTHODIR)

solve $\mathbf{u}$ from $\mathbf{K}\mathbf{u}=\mathbf{c}_{k-1}$,
$\mathbf{c}=\mathbf{A}\mathbf{u}$, orthogonalize

GMRES:
$\mathbf{y}_k$ solves $\min\{\|\mathbf{r}_0-\mathbb{A}\mathbf{y}\|\mid\mathbf{y}\in\mathcal{K}_k(\mathbb{A};\mathbf{r}_0)\}$
ORTHODIR (and GCR):
$\mathbf{s}_k$ solves $\min\{\|\mathbf{r}_0-\mathbf{s}\|\mid\mathbf{s}\in\mathbb{A}\mathcal{K}_k(\mathbb{A};\mathbf{r}_0)\}$

Keeping costs (work & storage) limited

Naive approaches

- Restart after ℓ steps (GCR(ℓ))
- Truncate:

  work with $\mathbf{u}_{k-\ell}, \ldots, \mathbf{u}_{k-1}$

  $\mathbf{c}_{k-\ell}, \ldots, \mathbf{c}_{k-1}$,

  orthogonalize only w.r.t. the last ℓ vectors.

– Lost of super-linear convergence;

may lead to stagnation

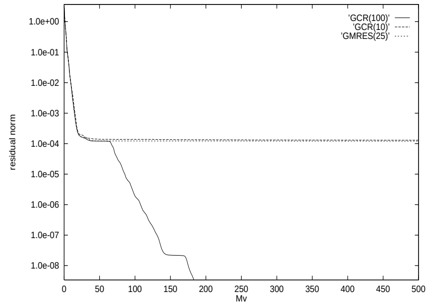
Advance approaches

Select
- implicitly or
- explicitly

an "interesting" part of the Krylov subspace

GCR(ℓ): restarted GCR

Choose $\mathbf{x}_0$
$k = 0$
$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$
while $\|\mathbf{r}_k\| > tol$ do
  solve $\mathbf{u}$ from $\mathbf{K}\,\mathbf{u} = \mathbf{r}_k$
  $\mathbf{c} \leftarrow \mathbf{A}\,\mathbf{u}$
  for $i = \lfloor \frac{k}{\ell} \rfloor \ell, \ldots, k-1$ do
    $\tau_{ik} = (c, \mathbf{c}_i)$, $\beta_{ik} = \tau_{ik}/\sigma_i$
    $\mathbf{u} \leftarrow \mathbf{u} - \beta_{ik}\,\mathbf{u}_i$
    $\mathbf{c} \leftarrow \mathbf{c} - \beta_{ik}\,\mathbf{c}_i$
  end for
  $\mathbf{u}_k = \mathbf{u}$, $\mathbf{c}_k = \mathbf{c}$
  $\sigma_k = (\mathbf{c}_k, \mathbf{c}_k)$, $\rho_k = (\mathbf{r}_k, \mathbf{c}_k)$, $\alpha_k = \rho_k/\sigma_k$
  $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\,\mathbf{u}_k$
  $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\,\mathbf{c}_k$
  $k \leftarrow k + 1$
end while

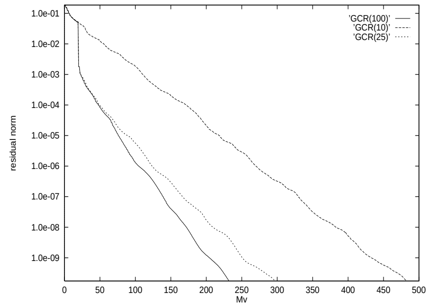


$-(Du_x)_x - (Du_y)_y = 1$ on $[0,1] \times [0,1]$.
$D = 1000$ for $0.1 \leq x, y \leq 0.9$ and $D = 1$ elsewhere.

Dirichlet BC on $y = 0$,
Neumann BC on other parts of Boundary.

82×82 discretization.

ILU Decomp. as preconditioner.



$u_{xx} + u_{yy} + u_{zz} + 1000 u_x = F$.

F is defined by the solution
$u(x,y,z) = xyz(1-x)(1-y)(1-z)$.

Discretized using $(22 \times 22 \times 22)$ volumes.
No preconditioning was used.

m–truncated GCR

Choose $\mathbf{x}_0$
$k = 0$
$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$
while $\|\mathbf{r}_k\| > tol$ do
  solve $\mathbf{u}$ from $\mathbf{K}\,\mathbf{u} = \mathbf{r}_k$
  $\mathbf{c} \leftarrow \mathbf{A}\,\mathbf{u}$
  for $i = \max(k-m, 0), \ldots, k-1$ do
    $\tau_{ik} = (c, \mathbf{c}_i)$, $\beta_{ik} = \tau_{ik}/\sigma_i$
    $\mathbf{u} \leftarrow \mathbf{u} - \beta_{ik}\,\mathbf{u}_i$
    $\mathbf{c} \leftarrow \mathbf{c} - \beta_{ik}\,\mathbf{c}_i$
  end for
  $\mathbf{u}_k = \mathbf{u}$, $\mathbf{c}_k = \mathbf{c}$
  $\sigma_k = (\mathbf{c}_k, \mathbf{c}_k)$, $\rho_k = (\mathbf{r}_k, \mathbf{c}_k)$, $\alpha_k = \rho_k/\sigma_k$
  $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\,\mathbf{u}_k$
  $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\,\mathbf{c}_k$
  $k \leftarrow k + 1$
end while

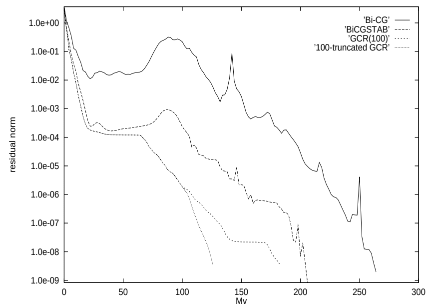
GMRES and GCR

In contrast to GCR,

GMRES and efficient GCR (ORTHODIR)

do **not** allow easy inclusion of

- truncation
- varying precondition



$-(Du_x)_x - (Du_y)_y = 1$ on $[0,1] \times [0,1]$.
$D = 1000$ for $0.1 \leq x, y \leq 0.9$ and $D = 1$ elsewhere.

Dirichlet BC on $y = 0$,
Neumann BC on other parts of Boundary.

82×82 discretization.

ILU Decomp. as preconditioner.

$$\min\{\|\mathbf{r}_0 - \mathbf{s}\| \mid \mathbf{s} \in \mathbb{A}\mathcal{V}\}$$

$\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{k-1}$ such that

$\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{k-1}$ is an orthogonal basis $\mathbb{A}\mathcal{V}$

where $\mathbf{c}_j = \mathbf{A}\,\mathbf{u}_j$ for $j = 0, 1, \ldots, k-1$

How to select the expansion vector for $\{\mathbf{c}_j\}$?

get $\angle(\mathbf{r}_0, \mathbb{A}\mathcal{V})$ small

Select a $\mathbf{u}$ (e.g., such that $\mathbf{u} \approx \mathbf{A}^{-1}\mathbf{r}_k$)

$$\mathbf{c}_k = \mathbf{A}\,\mathbf{u} - \sum_{j=0}^{k-1} \frac{(\mathbf{A}\mathbf{u}, \mathbf{c}_j)}{(\mathbf{c}_j, \mathbf{c}_j)}\mathbf{c}_j \quad \text{and}$$

$$\mathbf{u}_k = \mathbf{u} - \sum_{j=0}^{k-1} \frac{(\mathbf{A}\mathbf{u}, \mathbf{c}_j)}{(\mathbf{c}_j, \mathbf{c}_j)}\mathbf{u}_j$$

[van der Vorst & Vuik '93]

GCR$\star$

Choose $\mathbf{x}_0$
$k=0$
$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$
while $\|\mathbf{r}_k\| > tol$ do
  Select a $\mathbf{u}$ (e.g. s.t. $\mathbf{u} \approx \mathbf{A}^{-1}\mathbf{r}_k$)
  $\mathbf{c} \leftarrow \mathbf{A}\,\mathbf{u}$
  for $i = 0, \ldots, k-1$ do
    $\tau_{ik} = (\mathbf{c}, \mathbf{c}_i)$, $\beta_{ik} = \tau_{ik}/\sigma_i$
    $\mathbf{u} \leftarrow \mathbf{u} - \beta_{ik}\,\mathbf{u}_i$
    $\mathbf{c} \leftarrow \mathbf{c} - \beta_{ik}\,\mathbf{c}_i$
  end for
  $\mathbf{u}_k = \mathbf{u}$, $\mathbf{c}_k = \mathbf{c}$
  $\sigma_k = (\mathbf{c}_k, \mathbf{c}_k)$, $\rho_k = (\mathbf{r}_k, \mathbf{c}_k)$, $\alpha_k = \rho_k/\sigma_k$
  $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\,\mathbf{u}_k$
  $\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\,\mathbf{c}_k$
  $k \leftarrow k+1$
end while

Select a $\mathbf{u}$ (e.g., s.t. $\mathbf{u} \approx \mathbf{A}^{-1}\,\mathbf{r}_k$)

GMRESR

GMRESR(ℓ): solve $\mathbf{A}\,\mathbf{u} = \mathbf{r}_k$ approximately by ℓ steps of GMRES (To precondition, do it explicitly)

Other iterative methods in the inner-loop

- LSQR switch [van der Vorst Vuik '93]
- Bi-CGSTAB [de Sturler Fokkema '93]
- A polynomial method using a fixed polynomial p (for which $\mathbf{I} - \mathbf{A}p(\mathbf{A})$ is "small" and take $\mathbf{u} = p(\mathbf{A})\mathbf{r}_k$) (thus avoiding inner products)

Example: Chebychev iteration [van Gijzen '93]

Select a $\mathbf{u}$ (e.g., s.t. $\mathbf{u} \approx \mathbf{A}^{-1}\,\mathbf{r}_k$)

Deflating eigenvectors

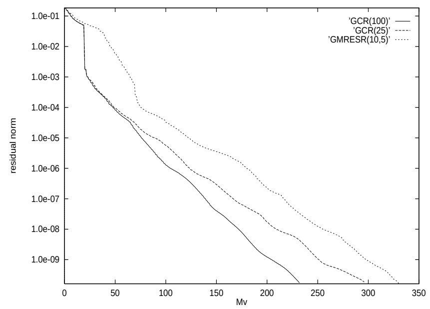
Suppose an (approximate) eigenvector $\tilde{\mathbf{u}}$ is explicitly known: $\mathbf{A}\tilde{\mathbf{u}} \approx \lambda\tilde{\mathbf{u}}$

Select in one (e.g. first) step $\mathbf{u} = \tilde{\mathbf{u}}$
select in other steps $\mathbf{u} \approx \mathbf{A}^{-1}\mathbf{r}_k$

In step "$\mathbf{u} = \tilde{\mathbf{u}}$" component of $\mathbf{x}$ (and $\mathbf{r}_0$) in the direction of $\tilde{\mathbf{u}}$ is computed. In subsequential steps, the search is in the space orthogonal to span($\tilde{\mathbf{u}}$) *(Deflation).*

Efficient for eigenvalues $\lambda \approx 0$,
no significant improvement if $\lambda \not\approx 0$.

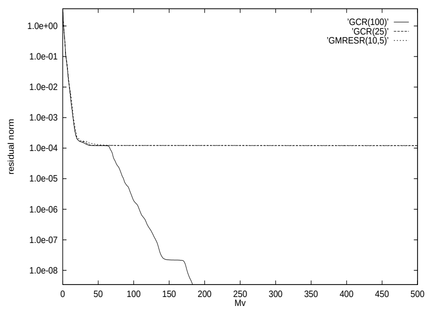
(Approximate) eigenvectors can be computed from iteration coeffients and vectors before, e.g., restart



$u_{xx} + u_{yy} + u_{zz} + 1000u_x = F.$

F is defined by the solution
$u(x, y, z) = xyz(1-x)(1-y)(1-z).$

Discretized using $(22 \times 22 \times 22)$ volumes.
No preconditioning was used.



$-(Du_x)_x - (Du_y)_y = 1$ on $[0,1] \times [0,1]$.
$D = 1000$ for $0.1 \le x, y \le 0.9$ and $D = 1$ elsewhere.

Dirichlet BC on $y = 0$,
Neumann BC on other parts of Boundary.
82×82 discretization.
ILU Decomp. as preconditioner.

Choosing ℓ in GMRESR(ℓ)

GMRES($m\ell$):
GMRES restarted every $m\ell$-th step.
GMRESR(m,ℓ):
GMRESR(ℓ) restarted every m-th step.

Experimentally: often GMRES($m\ell$) and GMRESR(m,ℓ) require the same number of MVs to obtain equally good approximations. One "sweep" requires $K := m\ell$ MV.

Per sweep:
AXPYs & DOTs GMRESR(m,ℓ) $\approx$ $(1/m + 2/\ell^2)\times$ # AXPYs & DOTs GMRES($m\ell$).

Optimal sweep length K depends on

- complexity per "step",
- accelerations

If optimal sweep length is K then choose ($m = K/\ell$ and) ℓ such that

$$m\,\ell = K \quad \text{minimizing} \quad 1/m + 2/\ell^2$$

Select a $\mathbf{u}$ (e.g., s.t. $\mathbf{A}\,\mathbf{u} = \mathbf{r}_k$)

$\mathbf{x}_k = \mathbf{x}_0 + \mathbf{y}_k$: the correction $\mathbf{y}_k$ belongs to span($\mathbf{u}_0, \ldots, \mathbf{u}_{k-1}$) and is optimal
in the sense that $\mathbf{r}_k \perp$ span($\mathbf{c}_0, \ldots, \mathbf{c}_{k-1}$).

Waste of computational effort to let the next $\mathbf{u}$ —the approximate solution of $\mathbf{A}\mathbf{u} = \mathbf{r}_k$— have components in span($\mathbf{u}_0, \ldots, \mathbf{u}_{k-1}$): $\mathbf{u}$ *should be such that $\mathbf{c} = \mathbf{A}\mathbf{u}$ is orthogonal to* span($\mathbf{c}_0, \ldots, \mathbf{c}_{k-1}$)

(Exact solution has this property.
For exact solution: $\mathbf{r}_{k+1} = \mathbf{r}_k - \mathbf{r}_k = 0$)

$\mathbf{P}_k = \sum_{i=0}^{k-1} \frac{\mathbf{c}_i\,\mathbf{c}_i^T}{\mathbf{c}_i^T\,\mathbf{c}_i}$ orthogonal projection onto span($\mathbf{c}_0, \ldots, \mathbf{c}_{k-1}$).

$\mathbf{Q}_k = \sum_{i=0}^{k-1} \frac{\mathbf{u}_i\,\mathbf{c}_i^T\mathbf{A}}{\mathbf{c}_i^T\,\mathbf{c}_i}$ skew projection onto span($\mathbf{u}_0, \ldots, \mathbf{u}_{k-1}$).

$$\mathbf{P}_k\mathbf{A} = \mathbf{A}\mathbf{Q}_k, \quad \mathbf{A}_k \equiv (\mathbf{I} - \mathbf{P}_k)\mathbf{A} = \mathbf{A}(\mathbf{I} - \mathbf{Q}_k)$$

$$\mathbf{A}_k\mathbf{u} = (\mathbf{I} - \mathbf{P}_k)\mathbf{A}\mathbf{u} = (\mathbf{I} - \mathbf{P}_k)\mathbf{r}_k = \mathbf{r}_k$$

GCRO$\star$:

Solve $\mathbf{u}$ approximately from $\mathbf{A}_k\,\mathbf{u} = \mathbf{r}_k$ such that $\mathbf{Q}_k\,\mathbf{u} = \mathbf{0}$.

Krylov subspace methods are very well suited:
if $\mathbf{Q}_k\,\mathbf{u}^{(0)} = \mathbf{0}$ then $\mathbf{Q}_k\,\mathbf{u}^{(j)} = \mathbf{0}$.

Note. $\mathbf{A}\,\mathbf{u}^{(\ell)}$ orthogonal $\mathbf{c}_0, \ldots, \mathbf{c}_{k-1}$:
no orthogonalisation needed in outer loop.

[de Sturler Fokkema '93]

GCRO⋆

Choose $\mathbf{x}_0$

$k = 0$

$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$

while $\|\mathbf{r}_k\| > tol$ do

> solve $\mathbf{u}$ approximately from
> $\mathbf{A}_k\,\mathbf{u} = \mathbf{r}_k$ such that $\mathbf{Q}_k\,\mathbf{u} = \mathbf{0}$

$\mathbf{u}_k = \mathbf{u}, \quad \mathbf{c}_k = \mathbf{A}\,\mathbf{u}$

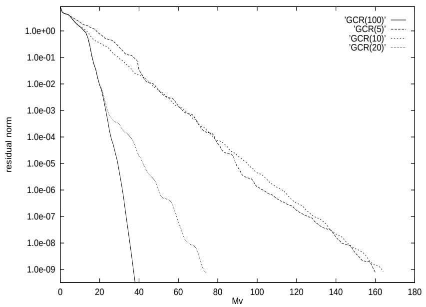
$\sigma_k = (\mathbf{c}_k, \mathbf{c}_k), \quad \rho_k = (\mathbf{r}_k, \mathbf{c}_k), \quad \alpha_k = \rho_k/\sigma_k$

$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\,\mathbf{u}_k$

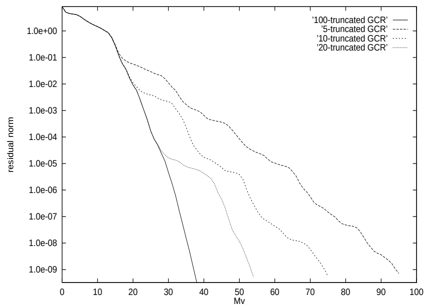
$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\,\mathbf{c}_k$

$k \leftarrow k+1$

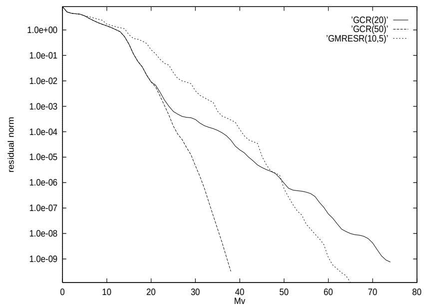
end while



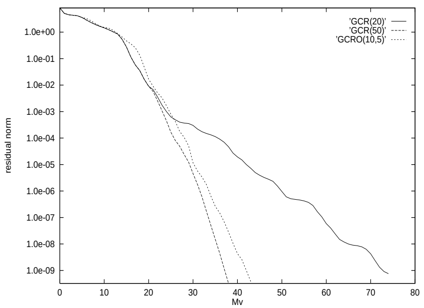
Sherman5 from the
Harwell-Boeing
Sparse Matrix Collection



Sherman5 from the
Harwell-Boeing
Sparse Matrix Collection



Sherman5 from the
Harwell-Boeing
Sparse Matrix Collection



Sherman5 from the
Harwell-Boeing
Sparse Matrix Collection

Other approaches

Implicit selection

There are other methods that allow varying preconditioners. They are not based on GCR.

- **F**lexible **GMRES** *based on GMRES* [Saad '93]

\+ A little more efficient than GMRESR,
\- no control on the residual

- **GEN**eralized **C**onjugate **G**radient [Axelsson '92]

Mathematically equivalent to GMRESR
\- inefficient (2MV per step)
\- less stable (uses classical Gram-Schmidt)

Explicit selection

Some methods try to identify converged eigenvector components and remove these eigenvectors from $\mathbf{A}$ by some deflation technique.

- Shift detected eigenvalues ≈ 0 to 1 [Kharchenko & Yeremin '94]
- Deflate eigenvector associated with eigenvalues ≈ 0 (take $\mathbf{u} =$ eigenvector) [Morgan '92]

A Flexible Arnoldi expansion

Expand recursively an orthonormal system $\mathbf{v}_1, \ldots, \mathbf{v}_k$ as follows

Select a $\mathbf{u}_k$, (e.g., $\mathbf{u}_k \approx \mathbf{A}^{-1}\mathbf{v}_k$)

$\tilde{\mathbf{v}} \leftarrow \mathbf{A}\,\mathbf{u}_k,\ \vec{h}_k = \mathbf{V}_k^*\,\tilde{\mathbf{v}},\ \tilde{\mathbf{v}} \leftarrow \tilde{\mathbf{v}} - \mathbf{V}_k\,\vec{h}_k$

$h_{k+1\,k} = \|\tilde{\mathbf{v}}\|,\ \mathbf{v}_{k+1} = \tilde{\mathbf{v}}/h_{k+1\,k}$

Theorem. Then, with $\mathbf{U}_k := [\mathbf{u}_1\,|\,\ldots\,|\,\mathbf{u}_k]$,

$$\boxed{\mathbf{A}\,\mathbf{U}_k = \mathbf{V}_{k+1}\,\underline{H}_k \quad \& \quad \mathbf{V}_{k+1}^*\mathbf{V}_{k+1} = I}$$

Examples.

- $\mathbf{u}_1 = \rho\,\mathbf{r}_0, \ldots, \mathbf{u}_k = \mathbf{v}_k$ (then $\mathbf{U}_k = \mathbf{V}_k$) [Arnoldi '51]
- $\mathbf{u}_1 = \mathbf{v}$ with $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{u}_1 = \rho\,\mathbf{r}_0, \ldots, \mathbf{u}_k = \mathbf{v}_k$ (deflation of eigenvectors) [Morgan '92]
- $\mathbf{u}_k$ solves $\mathbf{K}_k\,\mathbf{u} = \mathbf{v}_k$ [Saad '93]
- mixtures of the above

[Saad '93]

Flexible GMRES

With $\mathbf{r}_0 = \rho\mathbf{v}_1$, use Flexible Arnoldi to get

$$\mathbf{A}\mathbf{U}_k = \mathbf{V}_{k+1}\,\underline{H}_k \text{ with } \mathbf{V}_{k+1}^*\mathbf{V}_{k+1} = I$$

$$\begin{aligned} \|\,\mathbf{r}_k\,\| &= \min_{\vec{\gamma}} \|\,\mathbf{r}_0 - \mathbf{A}\mathbf{U}_k\vec{\gamma}\,\| \\ &= \min_{\vec{\gamma}} \|\,\mathbf{r}_0 - \mathbf{V}_{k+1}\,\underline{H}_k\,\vec{\gamma}\,\| \\ &= \min_{\vec{\gamma}} \|\,\rho\,e_1 - \underline{H}_k\,\vec{\gamma}\,\| \quad (\star) \end{aligned}$$

If $\vec{\gamma}$ solves $(\star)$ then

$$\mathbf{x}_k = \mathbf{x}_0 + \mathbf{U}_k\vec{\gamma}_k$$

If $\mathbf{K}_k = \mathbf{K}\ \forall k$ then $\mathbf{U}_k = \mathbf{K}^{-1}\mathbf{V}_k$ and

$$\mathbf{x}_k = \mathbf{x}_0 + \mathbf{K}^{-1}(\mathbf{V}_k\vec{\gamma}_k)$$

Suffices to store the $\mathbf{v}_j$