8 Conjugate Gradient method (CG)

Goal: iterative solution of $\mathbf{A}\mathbf{x}^* = \mathbf{b}$, $\mathbf{A} \in \mathbb{R}^{N\times N}$ symmetric positive definite
"rules": employ solely the matrix-vector multiplication $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$

Remark 8.1 *In many applications* $\mathbf{A}$ *can be very large but sparse, i.e.,* $\mathbf{A}$ *has only few non-zero entries per row. Then, the matrix-vector multiplication is feasible but a factorization of* $\mathbf{A}$ *may be infeasible (Cholesky factors of* $\mathbf{A}$ *typically need much more memory than* $\mathbf{A}$*).* ■

We will employ two scalar products:

- $(\mathbf{x},\mathbf{y})_2 := \mathbf{x}^T\mathbf{y} = \sum_i \mathbf{x}_i\mathbf{y}_i$ ("euklidean scalar product")
- $(\mathbf{x},\mathbf{y})_\mathbf{A} := \mathbf{x}^T\mathbf{A}\mathbf{y}$ ("energy scalar product")

Exercise 8.2 $(\cdot,\cdot)_\mathbf{A}$ *is a scalar product and* $\|\mathbf{x}\|_\mathbf{A} := \sqrt{(\mathbf{x},\mathbf{x})_\mathbf{A}}$ *is a norm on* $\mathbb{R}^N$.

Notation:

- $\mathbf{x}_0 \in \mathbb{R}^N$ arbitrary (=initial value)
- $\mathbf{x}^*$ solution of $\mathbf{A}\mathbf{x}^* = \mathbf{b}$
- $\mathbf{r}_0 := \mathbf{b} - \mathbf{A}\mathbf{x}_0$ = initial residual
- $\mathbf{e}_0 := \mathbf{x}^* - \mathbf{x}_0$ = initial error
- define, for each $\ell \in \mathbb{N}$ the *Krylov space*

$$\mathcal{K}_\ell := \mathcal{K}_\ell(\mathbf{A},\ \mathbf{r}_0) = \operatorname{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \ldots, \mathbf{A}^{\ell-1}\mathbf{r}_0\}$$

We have the *residual equation*

$$\mathbf{A}\mathbf{e}_0 = \mathbf{r}_0 \tag{8.1}$$

Question: Can one approximate $\mathbf{e}_0$ well from the spaces $\mathcal{K}_\ell$ (for "small" ℓ)? Consider the *best approximation*

$$\text{find } \tilde{\mathbf{e}}_\ell \in \mathcal{K}_\ell, \text{ s.t. } \|\mathbf{e}_0 - \tilde{\mathbf{e}}_\ell\|_\mathbf{A} \le \|\mathbf{e}_0 - \mathbf{x}\|_\mathbf{A} \qquad \forall \mathbf{x} \in \mathcal{K}_\ell \tag{8.2}$$

Correspondingly, one obtains an approximation $\mathbf{x}_\ell := \mathbf{x}_0 + \tilde{\mathbf{e}}_\ell$ of the original problem. Since $\mathbf{e}_0 = \mathbf{x}^* - \mathbf{x}_0$, we may characterize $\mathbf{e}_\ell$ also as:

$$\text{find } \mathbf{x}_\ell \in \mathbf{x}_0 + \mathcal{K}_\ell \text{ s.t. } \|\mathbf{x}^* - \mathbf{x}_\ell\|_\mathbf{A} \le \|\mathbf{x}^* - \mathbf{x}\|_\mathbf{A} \qquad \forall \mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_\ell \tag{8.3}$$

The solution $\mathbf{x}_\ell$ of (8.3) can also be characterized as follows:

Lemma 8.3 *The following are equivalent for* $\mathbf{x}_\ell \in \mathbf{x}_0 + \mathcal{K}_\ell$*:*

(i) $\mathbf{x}_\ell$ *solves (8.3)*

(ii) $(\mathbf{x}^* - \mathbf{x}_\ell, \mathbf{v})_\mathbf{A} = 0 \quad \forall \mathbf{v} \in \mathcal{K}_\ell$

(iii) $(\mathbf{r}_\ell, \mathbf{v})_2 = 0 \quad \forall \mathbf{v} \in \mathcal{K}_\ell$, *where* $\mathbf{r}_\ell := \mathbf{b} - \mathbf{A}\mathbf{x}_\ell$

Proof:
(ii) $\Leftrightarrow$ (iii): $(\mathbf{x}^* - \mathbf{x}_\ell, \mathbf{v})_\mathbf{A} = 0 \ \forall \mathbf{v} \in \mathcal{K}_\ell \Leftrightarrow (\mathbf{A}(\mathbf{x}^* - \mathbf{x}_\ell), \mathbf{v})_2 = 0 \ \forall \mathbf{v} \in \mathcal{K}_\ell \Leftrightarrow (\mathbf{b} - \mathbf{A}\mathbf{x}_\ell, \mathbf{v})_2 = 0 \ \forall \mathbf{v} \in \mathcal{K}_\ell$
(i) $\Longrightarrow$ (ii): Define for arbitrary $\mathbf{v} \in \mathcal{K}_\ell$ the function $\pi : \mathbb{R} \to \mathbb{R}$ by $\pi(t) := \|\mathbf{x}^* - \mathbf{x}_\ell + t\mathbf{v}\|_\mathbf{A}^2 = \|\mathbf{x}^* - \mathbf{x}_\ell\|_\mathbf{A}^2 + 2t(\mathbf{x}^* - \mathbf{x}_\ell,\, \mathbf{v})_\mathbf{A} + t^2|\mathbf{v}|_\mathbf{A}^2$
By assumption, π has a minimum at $t = 0 \Rightarrow 0 = \pi'(0) = (\mathbf{x}^* - \mathbf{x}_\ell, \mathbf{v})_\mathbf{A}$.
Since $\mathbf{v} \in \mathcal{K}_l$ is arbitrary, the claim follows.
(ii) $\Longrightarrow$ (i): Let $\mathbf{x}_\ell \in \mathbf{x}_0 + \mathcal{K}_\ell$ be such that $(\mathbf{x}^* - \mathbf{x}_\ell, \mathbf{v})_\mathbf{A} = 0 \ \forall \mathbf{v} \in \mathcal{K}_\ell$
Hence, the function π defined above has, for each fixed $\mathbf{v}$, its minimum at $t = 0 \Longrightarrow \pi(0) \leq \pi(t) \ \forall t \in \mathbb{R}$. Since $\mathbf{v}$ is arbitrary, (i) follows. □

For small ℓ, the $\ell \times \ell$ linear system of equations corresponding to (8.2) (or, alternatively, (8.3)) could be set up and solved (exercise!). However, the CG-algorithm proceeds in a much more economical way that determines $\mathbf{x}_\ell$ as a cheap update of $\mathbf{x}_{\ell-1}$.
$\mathbf{x}_\ell - \mathbf{x}_0 \in \mathcal{K}_\ell$ implies

$$\mathbf{r}_\ell = \mathbf{b} - \mathbf{A}\mathbf{x}_\ell = \mathbf{b} - \mathbf{A}\mathbf{x}_0 - \mathbf{A}(\mathbf{x}_\ell - \mathbf{x}_0) = \underbrace{\mathbf{r}_0}_{\in \mathcal{K}_0 \subset \mathcal{K}_{\ell+1}} \underbrace{- \mathbf{A}(\underbrace{\mathbf{x}_\ell - \mathbf{x}_0}_{\in \mathcal{K}_\ell})}_{\in \mathcal{K}_{\ell+1}} \in \mathcal{K}_{\ell+1},$$

i.e., $\mathbf{r}_\ell \in \mathcal{K}_{\ell+1}$. Since $\mathbf{r}_\ell$ is orthogonal to $\mathcal{K}_\ell$ (cf. Lemma 8.3,(iii)), we obtain inductively that

$$\mathcal{K}_{\ell+1} = \operatorname{span}\{\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_\ell\}.$$

We now focus on the algorithmic construction of the approximations $\mathbf{x}_\ell$. To that end, it is convenient to determine vectors $\mathbf{d}_0, \mathbf{d}_1, \ldots$, such that $\{\mathbf{d}_0, \ldots, \mathbf{d}_\ell\}$ is an orthogonal basis (w.r.t. the $(\cdot,\cdot)_\mathbf{A}$-scalar product) of $\mathcal{K}_{\ell+1}$. This is achieved with Gram-Schmidt orthogonalization: In view of $\mathcal{K}_\ell = \operatorname{span}\{\mathbf{r}_0, \ldots, \mathbf{r}_{\ell-1}\} = \operatorname{span}\{\mathbf{d}_0, \ldots, \mathbf{d}_{\ell-1}\}$ and $\mathcal{K}_{\ell+1} = \operatorname{span}\{\mathbf{r}_0, \ldots, \mathbf{r}_\ell\}$, we have that $\mathbf{d}_\ell$ has the form

$$\mathbf{d}_\ell = \mathbf{r}_\ell - \sum_{i=0}^{\ell-1} \beta_i \mathbf{d}_i$$

for suitable β_i. The orthogonality conditions

$$(\mathbf{d}_\ell, \mathbf{d}_i)_\mathbf{A} = 0 \text{ for } 0 \leq i \leq \ell - 1$$

produce

$$\beta_i = \frac{(\mathbf{r}_\ell, \mathbf{d}_i)_\mathbf{A}}{\|\mathbf{d}_i\|_\mathbf{A}^2}, \qquad i = 0, \ldots, \ell - 1.$$

For $i \leq \ell - 2$ we have

$$\beta_i = \frac{(\mathbf{r}_\ell, \mathbf{d}_i)_\mathbf{A}}{\|\mathbf{d}_i\|_\mathbf{A}^2} = \frac{(\mathbf{r}_\ell, \overbrace{\mathbf{A}\mathbf{d}_i}^{\in \mathbf{A}\mathcal{K}_{i+1} \subset \mathcal{K}_{i+2} \subset \mathcal{K}_\ell = \operatorname{span}\{\mathbf{r}_0, \ldots, \mathbf{r}_{\ell-1}\}})_2}{\|\mathbf{d}_i\|_\mathbf{A}^2} \overset{\textit{Lemma } 8.3,(iii)}{=} 0.$$

Therefore,

$$\mathbf{d}_\ell = \mathbf{r}_\ell - \beta_{\ell-1}\mathbf{d}_{\ell-1}, \quad \beta_{\ell-1} = \frac{(\mathbf{r}_\ell, \mathbf{d}_{\ell-1})_\mathbf{A}}{\|\mathbf{d}_{\ell-1}\|_\mathbf{A}^2}. \tag{8.4}$$

Next, we derive recursions for the $\mathbf{x}_\ell$ and $\mathbf{r}_\ell$: Since $\mathbf{x}_\ell - \mathbf{x}_{\ell-1} = (\mathbf{x}_\ell - \mathbf{x}_0) - (\mathbf{x}_0 - \mathbf{x}_{\ell-1}) \in \mathcal{K}_\ell$ and the orthogonality of Lemma 8.3,(ii) implies that $\mathbf{x}_\ell - \mathbf{x}_{\ell-1} = (\mathbf{x}_\ell - \mathbf{x}^*) - (\mathbf{x}^* - \mathbf{x}_{\ell-1})$ is $(\cdot,\cdot)_\mathbf{A}$-orthogonal to $\mathcal{K}_{\ell-1}$ we conclude

$$\mathbf{x}_\ell - \mathbf{x}_{\ell-1} = \alpha_\ell \mathbf{d}_{\ell-1}$$

for some $\alpha_\ell \in \mathbb{R}$. To derive an equation for the unknown α_ℓ we note that applying $\mathbf{A}$ to this equation yields

$$\alpha_\ell \mathbf{A}\mathbf{d}_{\ell-1} = \mathbf{A}(\mathbf{x}_\ell - \mathbf{x}_{\ell-1}) = \mathbf{A}\mathbf{x}_\ell - \mathbf{b} - (\mathbf{A}\mathbf{x}_{\ell-1} - \mathbf{b}) = -\mathbf{r}_\ell + \mathbf{r}_{\ell-1}$$

so that

$$\alpha_\ell(\mathbf{A}\mathbf{d}_{\ell-1}, \mathbf{r}_{\ell-1})_2 = (-\mathbf{r}_\ell + \mathbf{r}_{\ell-1}, \mathbf{r}_{\ell-1})_2 \overset{\text{Lemma 8.3, (iii)}}{=} \|\mathbf{r}_{\ell-1}\|_2^2. \tag{8.5}$$

We have thus obtained:

(α) $\mathbf{d}_\ell = \mathbf{r}_\ell - \beta_{\ell-1}\mathbf{d}_{\ell-1}$, $\beta_{\ell-1}$ given by (8.4)

(β) $\mathbf{r}_\ell = \mathbf{r}_{\ell-1} - \alpha_\ell \mathbf{A}\mathbf{d}_{\ell-1}$, α_ℓ given by (8.5).

(γ) $\mathbf{x}_\ell = \mathbf{x}_{\ell-1} + \alpha_\ell \mathbf{d}_{\ell-1}$

Remark 8.4 *Computationally, is it better to compute α_ℓ, β_ℓ as follows:*

$$\alpha_\ell = \frac{\|\mathbf{r}_{\ell-1}\|_2^2}{(\mathbf{d}_{\ell-1}, \mathbf{r}_{\ell-1})_\mathbf{A}} = \frac{\|\mathbf{r}_{\ell-1}\|_2^2}{(\mathbf{d}_{\ell-1}, \mathbf{d}_{\ell-1} + \beta_{\ell-2}\mathbf{d}_{\ell-2})_\mathbf{A}} = \frac{\|\mathbf{r}_{\ell-1}\|_2^2}{\|\mathbf{d}_{\ell-1}\|_\mathbf{A}^2}$$

$$\beta_{\ell-1} = \frac{(\mathbf{r}_\ell, \mathbf{d}_{\ell-1})_\mathbf{A}}{\|\mathbf{d}_{\ell-1}\|_\mathbf{A}^2} = \frac{(\mathbf{r}_\ell, \mathbf{A}\mathbf{d}_{\ell-1})_2}{\|\mathbf{d}_{\ell-1}\|_\mathbf{A}^2} = -\frac{(\mathbf{r}_\ell, \frac{\mathbf{r}_\ell - \mathbf{r}_{\ell-1}}{\alpha_\ell})_2}{\|\mathbf{d}_{\ell-1}\|_\mathbf{A}^2} = \frac{-\|\mathbf{r}_\ell\|_2^2}{\alpha_\ell\|\mathbf{d}_{\ell-1}\|_\mathbf{A}^2} = -\frac{\|\mathbf{r}_\ell\|_2^2}{\|\mathbf{r}_{\ell-1}\|_2^2}$$

We have thus derived the following algorithm:

Algorithm 8.5 (CG) *% input: SPD matrix $\mathbf{A}$, $\mathbf{b} \in \mathbb{R}^N$, initial vector $\mathbf{x}_0$*
% output: (approx.) solution $\mathbf{x}_n \approx \mathbf{A}^{-1}\mathbf{b}$

$\mathbf{r}_0 := \mathbf{b} - \mathbf{A}\mathbf{x}_0$, $\mathbf{d}_0 := \mathbf{r}_0$
for $\ell = 1, \ldots,$ *until stopping criterion is satisfied* **do** {
  $\alpha_\ell := \frac{\|\mathbf{r}_{\ell-1}\|_2^2}{\|\mathbf{d}_{\ell-1}\|_A^2}$
  $\mathbf{r}_\ell := \mathbf{r}_{\ell-1} - \alpha_\ell \mathbf{A}\mathbf{d}_{\ell-1}$
  $\mathbf{x}_\ell := \mathbf{x}_{\ell-1} + \alpha_\ell \mathbf{d}_{\ell-1}$
  $\beta_{\ell-1} := -\frac{\|\mathbf{r}_\ell\|_2^2}{\|\mathbf{r}_{\ell-1}\|_2^2}$
  $\mathbf{d}_\ell := \mathbf{r}_\ell - \beta_{\ell-1}\mathbf{d}_{\ell-1}$
}

slide 36

Remark 8.6 • *CG ist very economical w.r.t. memory requirements: merely 4 vectors of length N have to be kept in memory concurrently ($\mathbf{x}_\ell$, $\mathbf{r}_\ell$, $\mathbf{d}_\ell$, $\mathbf{A}\mathbf{d}_\ell$).*

- *In exact arithmetic, CG terminates with the exact solution after at most N steps. Technically, one may view CG therefore as a direct solver. Round-off problems, however, stop the method from realizing the exact solution after N steps.*

8.1 convergence behavior of CG

- $\mathbf{A} \in \mathbb{R}^{N\times N}$ SPD $\Rightarrow \exists$ *ONB* $\{\xi_1, \dots, \xi_N\}$ of $\mathbb{R}^N$ consisting of eigenvectors of $\mathbf{A}$ with corresponding eigenvalues $\lambda_i,\ i = 1,\dots,N$
- $\mathbf{x} = \sum_{i=1}^N \mathbf{x}_i\xi_i \Rightarrow \|\mathbf{x}\|_{\mathbf{A}}^2 = (\mathbf{x}, \mathbf{A}\mathbf{x})_2 = \sum_{i,j}(\mathbf{x}_i\xi_i, \lambda_j\mathbf{x}_j\xi_j) = \sum_{i,j}\mathbf{x}_i^2\lambda_j\delta_{ij} = \sum_{i=1}^N \mathbf{x}_i^2\lambda_i$
- $p \in \mathcal{P}_m \wedge \mathbf{x} = \sum_i \mathbf{x}_i\xi_i \Rightarrow p(\mathbf{A})\mathbf{x} = \sum_i p(\lambda_i)\xi_i\mathbf{x}_i$:
  Write $p(\mathbf{A}) = \sum_{j=0}^m p_j\mathbf{A}^j \Longrightarrow p(\mathbf{A})\mathbf{x} = \sum_j p_j\mathbf{A}^j\sum_i \mathbf{x}_i\xi_i = \sum_{i,j} p_j\mathbf{x}_i\lambda_i^j\xi_i = \sum_i \mathbf{x}_i\xi_i p(\mathbf{A}_i)$
- $p \in \mathcal{P}_m$ and $\mathbf{x} = \sum \mathbf{x}_i\xi_i \Longrightarrow \|p(\mathbf{A})\mathbf{x}\|_{\mathbf{A}}^2 = \sum \mathbf{x}_i^2|p(\lambda_i)|^2\lambda_i$:

$$\begin{aligned}
\|p(\mathbf{A})\mathbf{x}\|_{\mathbf{A}}^2 &= (p(\mathbf{A})\mathbf{x}, \mathbf{A}p(\mathbf{A})\mathbf{x})_2 = \sum_{i,j}(\mathbf{x}_i\xi_i p(\lambda_i), \mathbf{x}_j\xi_j\lambda_j p(\lambda_j))_2 \\
&= \sum_{i,j}\mathbf{x}_i\mathbf{x}_j p(\lambda_i)p(\lambda_j)\lambda_j \underbrace{(\xi_i,\xi_j)_2}_{\delta_{i,j}} = \sum_i |\mathbf{x}_i|^2\lambda_i|p(\lambda_i)|^2
\end{aligned}$$

In view of $\mathcal{K}_\ell = \operatorname{span}\{\mathbf{r}_0, \dots, \mathbf{A}^{\ell-1}\mathbf{r}_0\} = \{q(\mathbf{A})\mathbf{r}_0 \,|\, q \in \mathcal{P}_{\ell-1}\}$ and $\mathbf{r}_0 = \mathbf{A}\mathbf{e}_0$:

$$\begin{aligned}
\|\mathbf{x}^* - \mathbf{x}_\ell\|_{\mathbf{A}} &= \min_{\mathbf{x}\in\mathbf{x}_0+\mathcal{K}_\ell}\|\mathbf{x}^*-\mathbf{x}\|_{\mathbf{A}} = \min_{\mathbf{z}\in\mathcal{K}_\ell}\|\mathbf{e}_0-\mathbf{z}\|_{\mathbf{A}} = \min_{\mathbf{z}\in\mathcal{K}_\ell=\operatorname{span}\{\mathbf{r}_0,\dots,\mathbf{A}^{\ell-1}\mathbf{r}_0\}}\|\mathbf{e}_0-\mathbf{z}\|_{\mathbf{A}} \\
&= \min_{q\in\mathcal{P}_{\ell-1}}\|\mathbf{e}_0 - q(\mathbf{A})\mathbf{r}_0\|_{\mathbf{A}} = \min_{q\in\mathcal{P}_\ell:\,q(0)=0}\|\mathbf{e}_0 - q(\mathbf{A})\mathbf{e}_0\|_{\mathbf{A}} = \min_{q\in\mathcal{P}_\ell:\,q(0)=1}\|q(\mathbf{A})\mathbf{e}_0\|_{\mathbf{A}}.
\end{aligned}$$

Therefore:

Theorem 8.7 *The iterates $\mathbf{x}_\ell$ of the CG method satisfy*

$$\|\mathbf{x}^* - \mathbf{x}_\ell\|_{\mathbf{A}} = \min_{q\in\mathcal{P}_\ell:\,q(0)=1}\|q(\mathbf{A})\mathbf{e}_0\|_{\mathbf{A}}$$

We estimate further with $\mathbf{e}_0 = \sum \mathbf{x}_i\xi_i$:

$$\|q(\mathbf{A})\mathbf{e}_0\|_{\mathbf{A}}^2 \le \sum_i \mathbf{x}_i^2\lambda_i q^2(\lambda_i) \le \max_{\lambda\in\sigma(\mathbf{A})} q^2(\lambda)\sum_i \mathbf{x}_i^2\lambda_i = \max_{\lambda\in\sigma(\mathbf{A})} q^2(\lambda)\|\mathbf{e}_0\|_{\mathbf{A}}^2$$

Hence:

$$\|\mathbf{x}^* - \mathbf{x}_\ell\|_{\mathbf{A}} \le \min_{q\in\mathcal{P}_\ell:\,q(0)=1}\ \max_{\lambda\in\sigma(\mathbf{A})}|q(\lambda)|\ \|\mathbf{e}_0\|_{\mathbf{A}}$$

finis 14.DS

Theorem 8.8 *Let* $\mathbf{A} \in \mathbb{R}^{N\times N}$ *be SPD ,* $0 < \lambda_{min}(\mathbf{A}) \leq \lambda_{max}(\mathbf{A})$*,* $\kappa := \mathrm{cond}_2(\mathbf{A}) = \frac{\lambda_{max}(\mathbf{A})}{\lambda_{min}(\mathbf{A})}$*. Then: The iterates of the CG method satisfy*

$$\|\mathbf{x}^* - \mathbf{x}_\ell\|_{\mathbf{A}} \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^\ell \|\mathbf{e}_0\|_{\mathbf{A}}$$

Proof: We have

$$\|\mathbf{x}^* - \mathbf{x}_\ell\|_{\mathbf{A}} \leq \min_{q=\mathcal{P}_\ell,\, q(0)=1} \underbrace{\max_{\lambda\in\sigma(\mathbf{A})} |q(\lambda)|}_{\leq \max_{\mathbf{x}\in[\lambda_{min},\,\lambda_{max}]} |q(\mathbf{x})|} \|\mathbf{e}_0\|_{\mathbf{A}}$$

We select a specific q:

$$q(x) := \frac{T_\ell(\frac{a+b-2x}{b-a})}{T_\ell(\frac{b+a}{b-a})} \qquad a = \lambda_{min},\ b = \lambda_{max}$$

$T_\ell(x) =$ Chebyshev polynomial $= \frac{1}{2}\left[(t+\sqrt{t^2-1})^\ell + (t-\sqrt{t^2-1})^\ell\right]$ and uses

$$\max_{x\in[a,b]}\left|T_\ell\left(\frac{a+b-2x}{b-a}\right)\right| = \max_{x\in[-1,1]} |T_\ell(x)| = 1.$$

□

Remark 8.9 *Theorem 8.8 shows that the condition number of* $\mathbf{A}$ *is very important for the convergence behavior of the CG method. For matrices* $\mathbf{A}$ *with large condition number, one will therefore apply the CG not to* $\mathbf{A}$ *directly but to* $\mathbf{B}^{-1}\mathbf{A}$ *where the SPD matrix* $\mathbf{B}$ *is SPD. For more, see literature on the so-called "preconditioned CG" (PCG).* ■

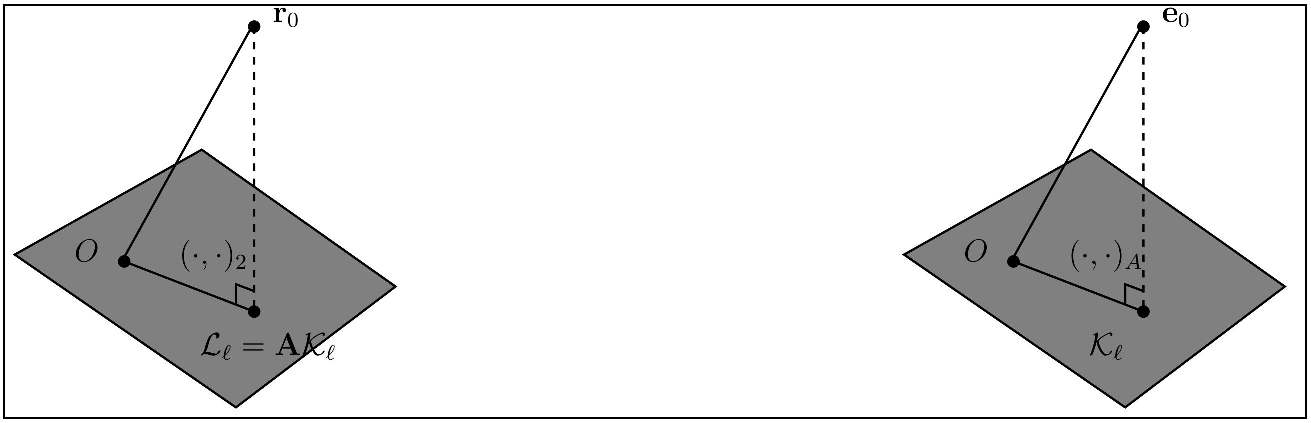

Figure 8.1: The orthogonality conditions (8.7) and Lemma 8.3, (ii).

8.2 GMRES (CSE)

goal: iterative methods for non-symmetric matrices $\mathbf{A} \in \mathbb{R}^{N\times N}$.
technique: for the Krylov space $\mathcal{K}_\ell := \text{span}\{\mathbf{r}_0, \dots, \mathbf{A}^{\ell-1}\mathbf{r}_0\}$ GMRES seek $\mathbf{x}_\ell \in \mathbf{x}_0 + \mathcal{K}_\ell$ such that

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}_\ell\|_2 \leq \|\mathbf{b} - \mathbf{A}x\|_2 \qquad \forall \mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_\ell \tag{8.6}$$

The mnimization property (8.6) implies an orthogonality condition:

Exercise 8.10 *Show that the residual $\mathbf{r}_\ell := \mathbf{b} - \mathbf{A}\mathbf{x}_\ell$ satisfies*

$$(\mathbf{b} - \mathbf{A}x_\ell)_2 = (\mathbf{r}_\ell, \mathbf{v})_2 = 0 \qquad \forall \mathbf{v} \in \mathbf{A}\mathcal{K}_\ell. \tag{8.7}$$

Hint: *Proceed as in the proof of Lemma 8.3 or in the derivation of the normal equations in Least Squares. (Note: GMRES can effectively be understood as a Least Squares method!)*

Remark 8.11 *The form (8.7) of GMRES suggests generalizations of GMRES: given a second space $\mathcal{L}_\ell$ one could consider: Find $\mathbf{x}_\ell \in \mathbf{x}_0 + \mathcal{K}_\ell$ such that*

$$(\mathbf{b} - \mathbf{A}x_\ell)_2 = (\mathbf{r}_\ell, \mathbf{v})_2 = 0 \qquad \forall \mathbf{v} \in \mathcal{L}_\ell. \tag{8.8}$$

Different choices of $\mathcal{L}_\ell$ lead to different method. The choice $\mathcal{L}_\ell = \mathcal{K}_\ell$ leads (for SPD matrices) to the CG-method (cf. Lemma 8.3), the choice $\mathcal{L}_\ell = \mathbf{A}\mathcal{K}_\ell$ to the classical GMRES. ■

One can show (this is not complicated, see literature) that for invertible matrices $\mathbf{A}$ GMRES finds (in exact arithmetic) the exact solution in N steps. As with the CG method, the importance lies in the fact that in practice good approximations are obtained $\ell \ll N$.

Computing the $\mathbf{x}_\ell$

As in the CG method, one computes the approximations $\mathbf{x}_\ell$ successively until one is found that is sufficiently accurate. It is, of course, essential that the $\mathbf{x}_\ell$ be computed efficiently from the orthogonality conditions (8.7). The general procedure is:

- Construct $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_\ell]$ an $N\times\ell$ matrix the columns of which form a basis for the space $\mathcal{K}_\ell$. It will be computationally convenient to choose the vectors $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ orthogonal.

- Construct $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_\ell]$ an $n \times \ell$ matrix the columns of which form a basis for the space $\mathcal{L}_\ell = \mathbf{A}\mathcal{K}_\ell$.
- Write the approximate solution as

  $$\mathbf{x}_\ell = \mathbf{x}_0 + \mathbf{V}\mathbf{y},$$

  where $\mathbf{y} \in \mathbb{R}^\ell$ is the vector of weights to be determined.
- Enforcing the orthogonality conditions (8.7) the system of equations

  $$\mathbf{W}^T\mathbf{A}\mathbf{V}\mathbf{y} = \mathbf{W}^T\mathbf{r}_0, \tag{8.9}$$

  from which the approximate solution $\mathbf{x}_\ell$ can be written as

  $$\mathbf{x}_\ell = \mathbf{x}_0 + \mathbf{V}(\mathbf{W}^T\mathbf{A}\mathbf{V})^{-1}\mathbf{W}^T\mathbf{r}_0. \tag{8.10}$$

  We note that the matrix $\mathbf{W}^T\mathbf{A}\mathbf{V}$ is only of the order $\ell \times \ell$; therefore its inversion is affordable (for $\ell \ll N$).

8.2.1 realization of the GMRES method

GMRES computes the vectors $\mathbf{v}_1, \ldots,$ successively such that the $\{\mathbf{v}_1, \ldots, \mathbf{v}_\ell\}$ is a basis of $\mathcal{K}_\ell$. Then the linear system described by (8.10) is solved in an efficient way.
We note (exercise!) that $\mathcal{K}_{\ell+1} = \mathbf{A}\mathcal{K}_\ell \supset \mathcal{K}_\ell$ for all ℓ. In the following, we will make the assumption that the inclusion is strict: $\mathcal{K}_\ell \subsetneqq \mathcal{K}_{\ell+1}$ for all ℓ of interest. That is, $\dim \mathcal{K}_\ell = \ell + 1$. One can show (see literature) that the case $\mathcal{K}_\ell = \mathcal{K}_{\ell+1}$ is a fortuitous case as then $\mathbf{x}_\ell = \mathbf{x}^*$ ("lucky breakdown").
The first step of the GMRES algorithm is to generate a vectors $\mathbf{v}_1, \ldots,$. Since we want the vectors $\mathbf{v}_j$, $j = 1, \ldots, \ell$, to be orthogonal, we will construct them using a variant of the Gram-Schmidt orthogonalization procedure given in Alg. 8.12 (in practice, a variant, the so-called "modified Gram-Schmidt" procedure, is used that is numerically more stable—see lines 5–8 of Alg. 8.14).

Algorithm 8.12 (Arnoldi, standard Gram-Schmidt variant)
% input $\mathbf{r}_0$*;*
%output: ONB of $\mathcal{K}_\ell = \mathrm{span}\{\mathbf{r}_0, \ldots, \mathbf{A}^{\ell-1}\mathbf{r}_0\}$

$\mathbf{v}_1 = \mathbf{r}_0/\|\mathbf{r}_0\|_2$
for $j = 1, 2, \ldots, \ell$ ***do***
  for $i = 1, 2, \ldots, j$ ***do***
    $h_{ij} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i)$
  end for
  $\mathbf{w}_j = \mathbf{A}\mathbf{v}_j - \sum_{i=1}^{j} h_{ij}\mathbf{v}_i$
  $h_{j+1,j} = \|\mathbf{w}_j\|_2$
  $\mathbf{v}_{j+1} = \mathbf{w}_j/h_{j+1,j}$
end for

The algorithm generates the $(\ell+1)\times\ell$ Hessenberg matrix

$$\bar{\mathbf{H}}_\ell = \begin{pmatrix} (\mathbf{A}\mathbf{v}_1,\mathbf{v}_1) & (\mathbf{A}\mathbf{v}_2,\mathbf{v}_1) & (\mathbf{A}\mathbf{v}_3,\mathbf{v}_1) & \dots & (\mathbf{A}\mathbf{v}_\ell,\mathbf{v}_1) \\ (\mathbf{A}\mathbf{v}_1,\mathbf{v}_2) & (\mathbf{A}\mathbf{v}_2,\mathbf{v}_2) & (\mathbf{A}\mathbf{v}_3,\mathbf{v}_2) & & \\ & (\mathbf{A}\mathbf{v}_2,\mathbf{v}_3) & (\mathbf{A}\mathbf{v}_3,\mathbf{v}_3) & & \\ & & (\mathbf{A}\mathbf{v}_3,\mathbf{v}_4) & \ddots & \\ & & & \ddots & (\mathbf{A}\mathbf{v}_\ell,\mathbf{v}_\ell) \\ & & & & (\mathbf{A}\mathbf{v}_\ell,\mathbf{v}_{\ell+1}) \end{pmatrix} \in \mathbb{R}^{(\ell+1)\times\ell}$$

together with the orthonormal vectors $\mathbf{v}_i = \frac{\mathbf{w}_{i-1}}{\|\mathbf{w}_{i-1}\|_2}$ that are produced by the Gram-Schmidt orthogonalization procedure:

$$\begin{aligned} \mathbf{w}_1 &= \mathbf{A}\mathbf{v}_1 - (\mathbf{A}\mathbf{v}_1,\mathbf{v}_1)\mathbf{v}_1 \\ \mathbf{w}_2 &= \mathbf{A}\mathbf{v}_2 - (\mathbf{A}\mathbf{v}_2,\mathbf{v}_1)\mathbf{v}_1 - (\mathbf{A}\mathbf{v}_2,\mathbf{v}_2)\mathbf{v}_2 \\ &\vdots \end{aligned}$$

as well as

$$\|\mathbf{w}_\ell\|_2 = (\mathbf{w}_\ell,\mathbf{v}_{\ell+1})_2 = (\mathbf{A}\mathbf{v}_\ell,\mathbf{v}_{\ell+1})_2 = h_{\ell+1,\ell}.$$

Exercise 8.13 *Assuming that Alg. 8.12 doesn't terminate prematurely, the vectors $\mathbf{v}_j$, $j = 1,\dots,\ell$, form an orthonormal basis of the Krylov space $\mathcal{K}_\ell$.*

We set $\mathbf{V}_\ell := (\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_\ell) \in \mathbb{R}^{N\times\ell}$. Since the vectors $\mathbf{v}_j$, $j = 1,\dots,\ell$, are orthonormal and since $\mathbf{A}\mathbf{v}_j \in \operatorname{span}\{\mathbf{v}_1,\dots,\mathbf{v}_{j+1}\}$ and thus $\mathbf{A}\mathbf{v}_j = \sum_{i=1}^{j+1}(\mathbf{A}\mathbf{v}_j,\mathbf{v}_i)\mathbf{v}_i$ we get

$$\begin{aligned} \mathbf{A}\mathbf{V}_\ell &= [\mathbf{A}\mathbf{v}_1 \quad \dots \quad \mathbf{A}\mathbf{v}_\ell] \\ &= \left[(\mathbf{A}\mathbf{v}_1,\mathbf{v}_1)\mathbf{v}_1 + (\mathbf{A}\mathbf{v}_1,\mathbf{v}_2)\mathbf{v}_2 \quad \dots \quad \sum_{i=1}^{\ell+1}(\mathbf{A}\mathbf{v}_\ell,\mathbf{v}_i)\mathbf{v}_i\right] \\ &= \mathbf{V}_{\ell+1}\bar{\mathbf{H}}_\ell. \end{aligned} \tag{8.11}$$

Additionally,

$$\mathbf{V}_\ell^\top \mathbf{A}\mathbf{V}_\ell = \mathbf{V}_\ell^\top \mathbf{V}_{\ell+1}\bar{\mathbf{H}}_\ell = [\,I\,|\,0\,]\bar{\mathbf{H}}_\ell = \mathbf{H}_\ell \tag{8.12}$$

where $\mathbf{H}_\ell$ is the square matrix obtained by removing the last row of $\bar{\mathbf{H}}_\ell$.
We abbreviate

$$\beta := \|\mathbf{r}_0\|_2$$

and note that $\beta\mathbf{v}_1 = \mathbf{r}_0$. Additionally, we observe $\beta\mathbf{V}_{\ell+1}\mathbf{e}_1 = \beta\mathbf{v}_1 = \mathbf{r}_0$, where $\mathbf{e}_1 = (1,0,0,\dots,0)^\top \in \mathbb{R}^{\ell+1}$.
GMRES minimizes the residuum (cf. (8.6)). Hence, seeking $\mathbf{x}_\ell$ in the form $\mathbf{x}_\ell = \mathbf{x}_0 + \mathbf{V}_\ell\mathbf{y}$ we can write

$$\begin{aligned} \mathbf{b} - \mathbf{A}\mathbf{x}_\ell &= \mathbf{b} - \mathbf{A}(\mathbf{x}_0 + \mathbf{V}_\ell\mathbf{y}) = \mathbf{r}_0 - \mathbf{A}\mathbf{V}_\ell\mathbf{y} = \beta\mathbf{v}_1 - \mathbf{V}_{\ell+1}\bar{\mathbf{H}}_\ell\mathbf{y} \\ &= \mathbf{V}_{\ell+1}(\beta\mathbf{e}_1 - \bar{\mathbf{H}}_\ell\mathbf{y}); \end{aligned}$$

exploiting the fact that the columns of $\mathbf{V}_{\ell+1}$ are orthonormal, we can determine the vector $\mathbf{y}$ by (8.6), i.e., $\mathbf{y}$ is the minimizer of

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}_\ell\|_2 = \min_{\mathbf{x}\in\mathbf{x}_0+\mathcal{K}_\ell} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 = \min_{\mathbf{y}\in\mathbb{R}^\ell} \|\beta\mathbf{e}_1 - \bar{\mathbf{H}}_\ell\mathbf{y}\|_2. \tag{8.13}$$

One way to solve for $\mathbf{y}$ is to set up and solve the *normal equations* using the Cholesky factorization with cost $O(\ell^3)$. However, since $\bar{\mathbf{H}}$ has Hessenberg form, its QR-factorization can be computed with $O(\ell^2)$ using, e.g., Givens rotations. The pseudo-code for the GMRES-algorithm now be given as Algorithm 8.14.

Algorithm 8.14 (GMRES (basic form)) *% input:* $\mathbf{x}_0$*, number of steps* ℓ

1: *Compute* $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$*,* $\beta = \|\mathbf{r}_0\|_2$*, and* $\mathbf{v}_1 = \mathbf{r}_0/\beta$
2: *Define the* $(\ell+1)\times\ell$ *matrix* $\bar{\mathbf{H}}_\ell$ *and set elements* h_{ij} *to zero*
3: ***for*** $j = 1, 2, \ldots, \ell$ ***do***
4:     $\mathbf{w}_j = \mathbf{A}\mathbf{v}_j$
5:     ***for*** $i = 1, \ldots, j$ ***do***
6:         $h_{ij} = (\mathbf{w}_j, \mathbf{v}_i)$
7:     ***end for***
8:     $\mathbf{w}_j = \mathbf{w}_j - h_{ij}\mathbf{v}_i$
9:     $h_{j+1,j} = \|\mathbf{w}_j\|_2$.
10:     **If** $h_{j+1,j} = 0$ **goto** *12*   *% lucky break—exact solution found*
11:     $\mathbf{v}_{j+1} = \mathbf{w}_j/h_{j+1,j}$
12: ***end for***
13: *Compute* $\mathbf{y}_\ell$ *as the minimizer of* $\|\beta\mathbf{e}_1 - \bar{\mathbf{H}}(1:j+1, 1:j)\mathbf{y}\|_2^2$ *(e.g., QR-factorization)*
14: $\mathbf{x}_\ell = \mathbf{x}_0 + \mathbf{V}_\ell\mathbf{y}_\ell$

A few comments concerning Alg. 8.14 are:

Remark 8.15

- The derivation of Alg. 8.14 assumed that matrix $\bar{\mathbf{H}}$ has full rank since we assumed that $\dim\mathcal{K}_\ell = \ell + 1$. Alg. 8.14 takes this into account by stopping if $h_{j+1,j} = 0$, which happens if $\mathcal{K}_{j+1} = \mathcal{K}_j$. However, a more careful analysis of the algorithm reveals that if $\bar{\mathbf{H}}$ does not have full rank, i.e., if $\mathcal{K}_j = \mathcal{K}_{j+1}$, then GMRES has actually found the exact solution $\mathbf{x}^*$. This situation is therefore called a "lucky breakdown".

- Solving the minimization problem in line 12 is done by QR-factorization of the Hessenberg matrix $\bar{\mathbf{H}}$, e.g., with Givens rotations.

- The algorithm is implemented differently in practice. The parameter ℓ is not determined *a priori*. Instead, a maximum number ℓ_{max} is given (typically dictated by the computational resources). The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_\ell$ are computed successively together with the matrices $\bar{\mathbf{H}}_\ell$; that is, if the vectors $\mathbf{v}_j$, $1 \leq j \leq \ell - 1$ and the matrix $\bar{\mathbf{H}}_{\ell-1}$ have already been computed, one merely needs to compute $\mathbf{v}_\ell$ and the matrix $\bar{\mathbf{H}}_\ell$ is obtained from $\bar{\mathbf{H}}_{\ell-1}$ by adding one column and the entry $h_{\ell+1,\ell}$. An appropriate termination condition (typically, the size of the residual $\|\mathbf{b} - \mathbf{A}\mathbf{x}_\ell\|_2$) is employed to stop the iteration. If the maximum number of iterations has been reached without triggering the termination condition, then a *restart* is done, i.e., GMRES is started afresh with the last approximation $\mathbf{x}_{\ell_{max}}$ as the initial guess. This is called *restarted GMRES*(ℓ_{max}) in the literature.

■

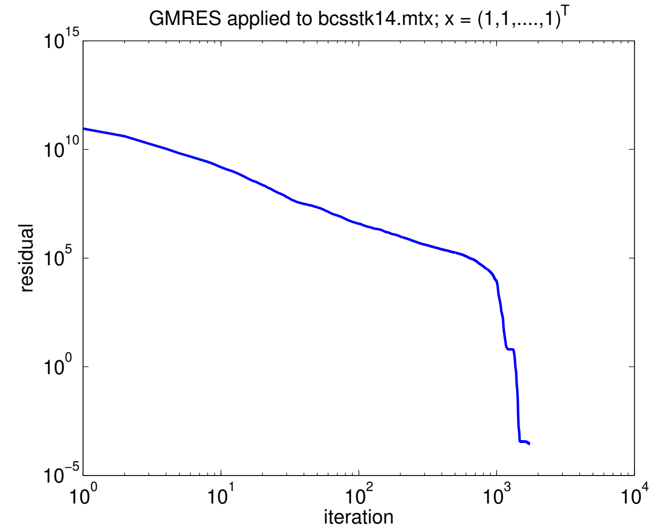

Figure 8.2: Convergence history of GMRES (A is SPD).

slide 37

Remark 8.16 *Faute de mieux*, the residual $\|\mathbf{b}-\mathbf{A}\mathbf{x}_\ell\|_2$ is typically used as a stopping criterion in GMRES. It should be noted that for matrices A with large $\kappa_2(\mathbf{A})$, the error may be large in spite of the residual being small:

$$\frac{\|\mathbf{x}-\mathbf{x}_\ell\|_2}{\|\mathbf{x}\|_2} \leq \kappa_2(\mathbf{A}) \frac{\|\mathbf{b}-\mathbf{A}\mathbf{x}_\ell\|_2}{\|\mathbf{b}\|_2}.$$

■

Example 8.17 MATLAB has a robust version of restarted GMRES that can be used for experimentation. Applying this version of GMRES to the SPD matrix $\mathbf{A} \in \mathbb{R}^{1806\times 1806}$ `bcsstk14.mtx` of `MatrixMarket` with exact solution $x = (1, 1, \cdots, 1)^\top$ results in the convergence history plotted in Fig. 8.2. We note that the residual decays as the number if iterations increases. If the number of iterations reaches the problem size, the exact solution should be found. As in this example, this doesn't happen in practice due to round-off problems, but the residual is quite small. It should be noted that, generally speaking, GMRES is employed in connection with a suitable preconditioner. We expect this to greatly improve the convergence behavior. ■