

Krylov-Unterraum-Approximationen an $\varphi_\ell(-tA)v$

Lena Martin

KARLSRUHER INSTITUT FÜR TECHNOLOGIE (KIT)



- Polynomialer Krylov-Unterraum:

$$\begin{aligned}\mathcal{K}_m(A, v) &= \text{span}\{v, Av, \dots, A^{m-1}v\} \\ &= \{\rho_{m-1}(A)v \mid \rho_{m-1} \in \mathcal{P}_{m-1}\}\end{aligned}$$

- Shift-and-Invert-Krylov-Unterräume:

$$\begin{aligned}\mathcal{K}_m((I + \gamma A)^{-1}, v) &= \text{span}\{v, (I + \gamma A)^{-1}v, \dots, (I + \gamma A)^{-(m-1)}v\} \\ &= \{\rho_{m-1}((I + \gamma A)^{-1})v, \quad \rho_{m-1} \in \mathcal{P}_{m-1}\} \\ &= \{r_{m-1}(A)v = \frac{\rho_{m-1}(A)}{(I + \gamma A)^{m-1}}v, \quad \rho_{m-1} \in \mathcal{P}_{m-1}\}\end{aligned}$$

- Rationale Krylov-Unterräume
- Erweiterte Krylov-Unterräume

Problemstellung: $\varphi_\ell(-tA)v$

- $A \in \mathbb{C}^{n \times n}$ dünn besetzt, $\mathbb{F}(A) \subseteq \mathbb{C}_0^+$
- A aus Ortsdiskretisierung zeitabhängiger PDE \rightarrow ODE

$$u'(t) = -Au(t) + \sum_{\ell=1}^k \alpha_\ell \frac{t^{\ell-1}}{(\ell-1)!} v, \quad u(0) = u_0$$

- Variation der Konstanten

$$u(t) = e^{-tA} u_0 + \sum_{\ell=1}^k \alpha_\ell t^\ell \varphi_\ell(-tA) v,$$

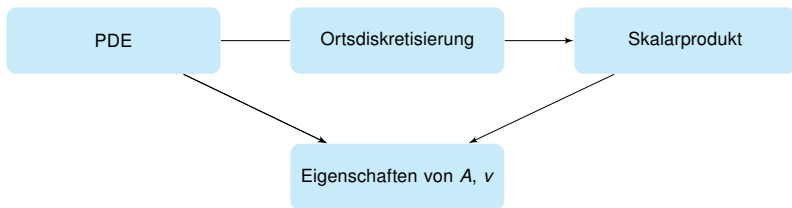
$$\varphi_0(z) = e^z, \quad \varphi_\ell(z) = \int_0^1 e^{(1-s)z} \frac{s^{\ell-1}}{(\ell-1)!} ds, \quad \ell \geq 1$$

- 1 Konstruktion einer ONB $\{v_1, \dots, v_m\}$ von $\mathcal{K}_m(A, v)$ oder $\mathcal{K}_m((I + \gamma A)^{-1}, v)$
- 2 $S_m = V_m^H A V_m \in \mathbb{C}^{m \times m}$
- 3 Approximation

$$\varphi_\ell(-tA)v \approx \|v\| V_m \varphi_\ell(-tS_m) e_1, \quad m \ll n$$

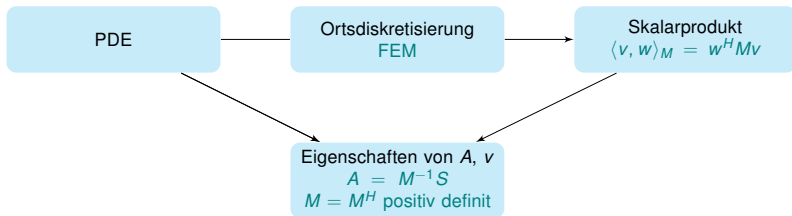
- 1 Konstruktion einer ONB $\{v_1, \dots, v_m\}$ von $\mathcal{K}_m(A, v)$ oder $\mathcal{K}_m((I + \gamma A)^{-1}, v)$
- 2 $S_m = V_m^H A V_m \in \mathbb{C}^{m \times m}$
- 3 Approximation

$$\varphi_\ell(-tA)v \approx \|v\| V_m \varphi_\ell(-tS_m) e_1, \quad m \ll n$$



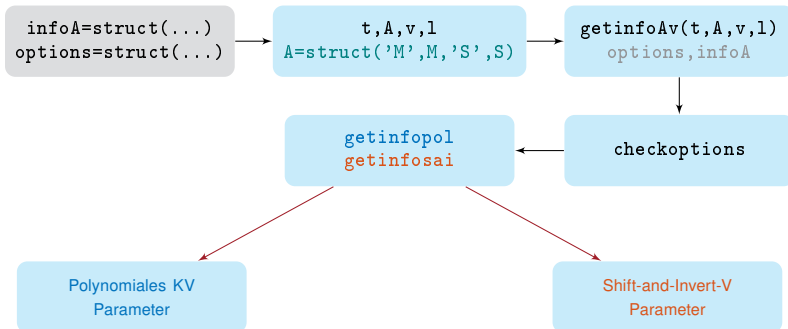
- 1 Konstruktion einer ONB $\{v_1, \dots, v_m\}$ von $\mathcal{K}_m(A, v)$ oder $\mathcal{K}_m((I + \gamma A)^{-1}, v)$
- 2 $S_m = V_m^H M M^{-1} S V_m = V_m^H S V_m, \quad V_m^H M V_m = I_m$
- 3 Approximation

$$\varphi_\ell(-tA)v \approx \|v\|_M V_m \varphi_\ell(-tS_m) e_1, \quad m \ll n$$

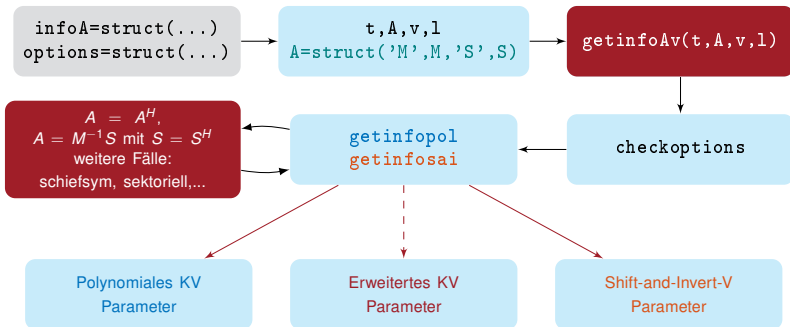


- Auswahl und Anwendung von KUV zur Approximation von $\varphi_\ell(-tA)v$ erleichtern
- Strategien entwickeln, Implementierung von Hilfsroutinen

- Auswahl und Anwendung von KUV zur Approximation von $\varphi_\ell(-tA)v$ erleichtern
- Strategien entwickeln, Implementierung von Hilfsroutinen



- Auswahl und Anwendung von KUV zur Approximation von $\varphi_\ell(-tA)v$ erleichtern
- Strategien entwickeln, Implementierung von Hilfsroutinen



Polynomial

Input : t, S, v, tol

Output : $u_m \approx \varphi_\ell(-tS)v$

```
1  $\beta := \|v\|, v_1 := \frac{v}{\beta}$ 
2
3 while Fehlerschätzer > tol do
4    $w := Sv_m$ 
5
6   for  $k=1:m$  do
7      $h_{k,m} = v_k^H w$ 
8      $v_{m+1} := v_{m+1} - h_{k,m}v_k$ 
9   end
10   $h_{m+1,m} := \|v_{m+1}\|$ 
11   $v_{m+1} := v_{m+1}/h_{m+1,m}$ 
12
13   $y_m := \varphi_\ell(-tH_m)e_1$ 
14
15   $e_m :=$  Fehlerschätzer
16 end
17  $u_m := \beta V_m y_m$ 
```

Shift-and-Invert

$w := v_1$

Löse LGS $(I + \gamma S)v_{m+1} = w$

$w := v_{m+1}$

$y_m := \varphi_\ell(-\frac{t}{\gamma}(I_m - H_m^{-1}))e_1$

Polynomial

Input : t, M, S, v, tol

Output : $u_m \approx \varphi_\ell(-tM^{-1}S)v$

```
1  $\beta := \|v\|_M, v_1 := \frac{v}{\beta}$ 
2 while Fehlerschätzer > tol do
3    $w := Sv_m$ 
4
5   Löse LGS  $Mv_{m+1} = w$ 
6
7   for  $k=1:m$  do
8      $h_{k,m} = v_k^H w$ 
9      $v_{m+1} := v_{m+1} - h_{k,m} v_k$ 
10  end
11
12   $h_{m+1,m} := \|v_{m+1}\|_M$ 
13
14   $v_{m+1} := v_{m+1} / h_{m+1,m}$ 
15   $y_m := \varphi_\ell(-tH_m)e_1$ 
16
17   $e_m :=$  Fehlerschätzer
18 end
19  $u_m := \beta v_m y_m$ 
```

Shift-and-Invert

$$w := Mv_1$$

Löse LGS $(M + \gamma S)v_{m+1} = w$

$$w := Mv_{m+1}$$

$$w := Mv_{m+1}$$

$$h_{m+1,m} := \sqrt{v^H w}$$

$$y_m := \varphi_\ell\left(-\frac{t}{\gamma}(I_m - H_m^{-1})\right)e_1$$

$$(I + \gamma A)^{-1} = (I + \gamma M^{-1}S)^{-1}v = (M + \gamma S)^{-1}Mv$$

Schätzen der benötigten Iterationen

a-priori Fehlerschranken abhängig von $\mathbb{F}(A)$

a-priori Fehlerschranken abhängig von $\mathbb{F}(A)$

M-Wertebereich

- $\varrho_A^M(x) = \frac{\langle Ax, x \rangle_M}{\langle x, x \rangle_M} = \frac{x^H S x}{x^H M x}, x \in \mathbb{C} \setminus \{0\}$

- $\mathbb{F}(S) \subset \mathbb{C}_0^+ \Rightarrow \mathbb{F}^M(A) \subset \mathbb{C}_0^+$

- $M^{-1}S$ ähnlich $M^{-1/2}SM^{-1/2}$

$$\begin{aligned} \Rightarrow \quad & \text{a) } \sigma(M^{-1}S) = \sigma(M^{-1/2}SM^{-1/2}) \\ & \text{b) } \mathbb{F}^M(A) = \mathbb{F}(M^{-1/2}SM^{-1/2}) \end{aligned}$$

Near-Best-Eigenschaft der M -Krylov-Unterraum-Approximationen

- Approximation:

$$u(t) = \varphi_\ell(-tA)v \approx \underbrace{\|v\|_M \varphi_\ell(-tV_m^H S V_m)}_{=: u_m(t)} e_1$$

- Fehler: $\epsilon_m(t) = u(t) - u_m(t)$
- Abschätzung für polynomiale Krylov-Unterraum-Approximation:

$$\|\epsilon_m(t)\|_M \leq 2C \|v\|_M \inf_{p \in \mathcal{P}^{m-1}} \max_{z \in \mathbb{F}^M(tA)} |\varphi_\ell(-z) - p(z)|.$$

Near-Best-Eigenschaft der M -Krylov-Unterraum-Approximationen

- Approximation:

$$u(t) = \varphi_\ell(-tA)v \approx \underbrace{\|v\|_M \varphi_\ell(-tV_m^H S V_m)}_{=: u_m(t)} e_1$$

- Fehler: $\epsilon_m(t) = u(t) - u_m(t)$
- Abschätzung für polynomiale Krylov-Unterraum-Approximation:

$$\|\epsilon_m(t)\|_M \leq 2C \|v\|_M \inf_{p \in \mathcal{P}^{m-1}} \max_{z \in \mathbb{F}^M(tA)} |\varphi_\ell(-z) - p(z)|.$$

- Abschätzung für Shift-and-Invert-Approximation:

$$\begin{aligned} \|\epsilon_m(t)\|_M &\leq 2C \|v\|_M \inf_{r \in \mathcal{R}_{m-1}^{m-1}(\gamma)} \max_{z \in \mathbb{F}^M(tA)} |\varphi_\ell(-z) - r(z)| \\ &= 2C \|v\|_M \inf_{p \in \mathcal{P}^{m-1}(\gamma)} \max_{z \in \mathbb{F}^M(A)} |f_\gamma^t(z) - p(z)| \end{aligned}$$

A-priori Fehlerabschätzungen für $S = S^H$

Polynomiale Krylov-Unterraum-Approximation an $\varphi_\ell(-tA)v$, $\ell \geq 0$

- Hochbruck, Lubich (1997, 2008, 2013)
- $\ell \geq 0$: superlineare Konvergenz für $m > \omega$

$$S = S^H$$

$$\omega = \sqrt{t \frac{\lambda_{\max}}{2}}$$

$$S = -S^H$$

$$\omega = t\lambda_{\max}$$

- $\lambda_{\max} = \max_{\lambda \in \sigma(A)} |\lambda|$

Spektralradius λ_{\max} schätzen

- M diagonal: Satz von Gershgorin
- Für verallgemeinertes EWP:

$$M^{-1}Sx = \lambda x, \quad x \neq 0 \quad \Leftrightarrow \quad Sx = \lambda Mx, \quad x \neq 0$$

- M diagonal: Satz von Gershgorin
- Für verallgemeinertes EWP:

$$M^{-1}Sx = \lambda x, \quad x \neq 0 \quad \Leftrightarrow \quad Sx = \lambda Mx, \quad x \neq 0$$

- Varga, Kostić, Cvetkovic (2009): verallgemeinerte Gershgorin-Mengen
- Nakatsukasa (2011): verallgemeinerte Gershgorin Kreise
- M strikt diagonaldominant (sdd), falls

$$\sum_{\substack{j=1 \\ j \neq i}}^n |m_{ij}| < |m_{ii}|, \quad i = 1, \dots, n$$

Verallgemeinerte Gershgorin Kreise

Einschlussmenge für verallgemeinerte Eigenwerte:

$$K(S, M) = \bigcup_{i=1}^n K_i(S, M), \quad \text{wobei } K_i(S, M) = K_i^M \cap K_i^S$$

Betrachte ***i*-te Zeile**

M

sdd: K_i^M Kreisscheibe um $\frac{s_{ii}}{m_{ii}}$

nsdd: $K_i^M = \mathbb{C}$

Einschlussmenge für verallgemeinerte Eigenwerte:

$$K(S, M) = \bigcup_{i=1}^n K_i(S, M), \quad \text{wobei } K_i(S, M) = K_i^M \cap K_i^S$$

Betrachte ***i*-te Zeile**

M

S

sdd: K_i^M Kreisscheibe um $\frac{s_{ii}}{m_{ii}}$

a) K_i^S Kreisscheibe

b) K_i^S äußeres einer Kreisscheibe

c) K_i^S Halbebene

In allen Fällen gilt $\frac{s_{ii}}{m_{ii}} \in K_i^S$.

nsdd: $K_i^M = \mathbb{C}$

$K_i^S = \mathbb{C}$

FEM-Diskretisierung hom. WLG

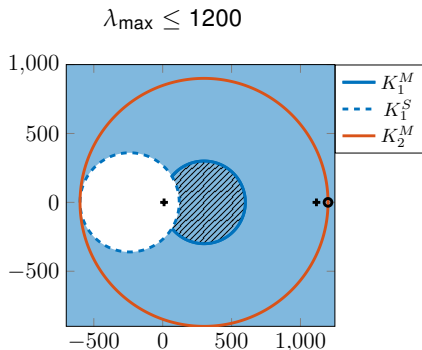
$$u_t(x, t) = u_{xx}(x, t) \quad \forall t > 0, x \in (0, 1), \quad u(0, t) = u(1, t) = 0, \quad u(x, 0) = u_0$$

- äquidistantes Gitter mit $n + 2$ Punkten, $h = \frac{1}{n+1} = \frac{1}{10}$

$$M = \frac{h}{6} \text{tridiag}(1, 4, 1)$$

$$S = \frac{1}{h} \text{tridiag}(-1, 2, -1)$$

$$\lambda_{\max} = \frac{6 \left(1 - \cos\left(\frac{n\pi}{n+1}\right)\right)}{h^2 \left(2 + \left(\frac{n\pi}{n+1}\right)\right)} \approx 1160$$



Alternativen?

- eigs zu teuer
- Sollte ohne Berechnung von $M^{-1}S$ berechenbar sein
- Kann vom Nutzer übergeben werden
- Erweitertes Krylov-Unterraum-Verfahren

Shift-and-Invert

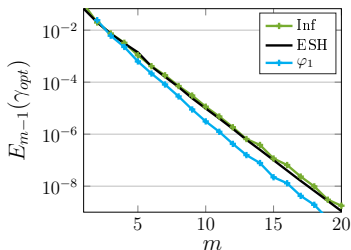
- Schranken abhängig von γ , asymptotisch: $\gamma = \frac{\sqrt{2}}{m}$
- Eshof, Hochbruck (2006): hermitesche Matrizen
- $e^{-tM^{-1}S}v, S = S^H$

$$\|\epsilon_m\|_M \leq 2\|v\|_M E_{m-1}(\tilde{\gamma}), \quad \tilde{\gamma} = \frac{\gamma}{t}$$

$$\begin{aligned} E_{m-1}(\gamma) &:= \inf_{r \in \mathcal{R}_{m-1}^{m-1}} \sup_{z \in [0, \infty)} |r(z) - \exp(-tz)| \\ &= \inf_{p \in \mathcal{P}_{m-1}} \sup_{w \in [0, 1]} |p(w) - f_\gamma^t(w)| \end{aligned}$$

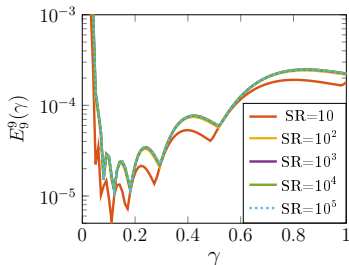
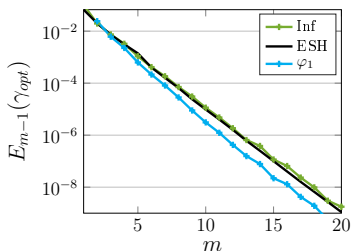
A-priori-Fehlerschranke für SaIA

- asymptotisch: $\gamma = \sqrt{2}/m$
- polynomiale Bestapproximation: Remez-Algorithmus
- Minimierung von $E_{m-1}(\gamma_{opt}(m, \lambda_{\max}))$



A-priori-Fehlerschranke für SaIA

- asymptotisch: $\gamma = \sqrt{2}/m$
- polynomiale Bestapproximation: Remez-Algorithmus
- Minimierung von $E_{m-1}(\gamma_{opt}(m, \lambda_{max}))$

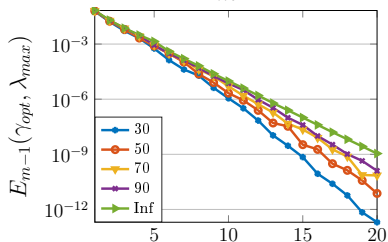
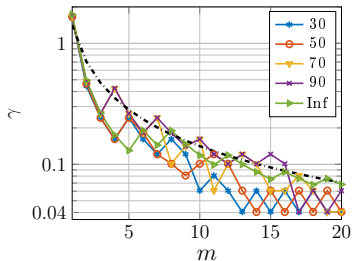


- Tabellen für φ_ℓ -Funktionen $\ell = 1, 2$

Wahl des Shift-Parameters γ

Mit Remez-Algorithmus ermittelte Werte für $\gamma_{opt}(m, \lambda_{max})$

λ_{max}	30	50	70	90	∞
$m = 1$	1.6162	1.6566	1.6768	1.6970	1.6970
$m = 2$	0.4444	0.4646	0.4848	0.4848	0.4848
$m = 3$	0.2424	0.2424	0.2626	0.2626	0.2626
$m = 4$	0.1616	0.1616	0.4242	0.4242	0.4242
$m = 5$	0.2424	0.2424	0.2626	0.2626	0.2626
$m = 6$	0.1616	0.1818	0.1818	0.1818	0.1818
$m = 7$	0.1212	0.1212	0.2424	0.2424	0.2424
$m = 8$	0.1616	0.1010	0.1010	0.1818	0.1818
$m = 9$	0.1212	0.0808	0.1414	0.1414	0.1414
$m = 10$	0.0606	0.1010	0.1616	0.1616	0.1616
$m = 11$	0.0808	0.1212	0.0606	0.1212	0.1212
$m = 12$	0.0606	0.1010	0.1010	0.1010	0.1010
$m = 13$	0.0404	0.0606	0.1212	0.1212	0.1212
$m = 14$	0.0606	0.0404	0.1010	0.1010	0.1010
$m = 15$	0.0404	0.0606	0.0606	0.1212	0.1212
$m = 16$	0.0606	0.0404	0.0606	0.1010	0.1010
$m = 17$	0.0404	0.0606	0.0808	0.0404	0.0404
$m = 18$	0.0606	0.0404	0.0606	0.0606	0.0606
$m = 19$	0.0404	0.0606	0.0404	0.0404	0.0404
$m = 20$	0.0404	0.0404	0.0404	0.0606	0.0606



`getinfoAv`

- prüft Symmetrieeigenschaften für S
- Schätzt, sofern möglich Spektralradius
- Berechnet $\|v\|_M$
- `options.tol`, default $1e-8$

→ entsprechende Schranken bzw. Schätzung

`[mMax, ω , bound, memoryv]=getinfoPol(t, infoA, v, l, options)`

`[mMax, γ , error, bound]=getinfoSai(t, infoA, v, l, options)`



J. M. Anderson und J. Clunie. "Isomorphisms of the disc algebra and inverse faber sets". *Mathematische Zeitschrift* 188.4 (1985), S. 545–558. DOI: [10.1007/BF01161656](https://doi.org/10.1007/BF01161656).



M. A. Botchev, V. Grimm und M. Hochbruck. "Residual, restarting and Richardson iteration for the matrix exponential". *SIAM J. Sci. Comput.* 35.3 (2013), A1376–A1397. DOI: [10.1137/110820191](https://doi.org/10.1137/110820191).



Bernhard Beckermann und Lothar Reichel. "Error Estimates and Evaluation of Matrix Functions via the Faber Transform". *SIAM J. Numerical Analysis* (2009), S. 3849–3883.



G.F. Carey und K. Sepehrnoori. "Gershgorin theory for stiffness and stability of evolution systems and convection-diffusion". *Computer Methods in Applied Mechanics and Engineering* 22.1 (1980), S. 23–48. DOI: [http://dx.doi.org/10.1016/0045-7825\(80\)90049-3](http://dx.doi.org/10.1016/0045-7825(80)90049-3).



J. van den Eshof und M. Hochbruck. "Preconditioning Lanczos approximations to the matrix exponential". *SIAM J. Sci. Comp.* 27.4 (2006), S. 1438–1457.



V. Kostić, R. S. Varga und L. Cvetković. "Localization of generalized eigenvalues by Cartesian ovals". *Numer. Linear Algebra Appl.* 19.4 (2012), S. 728–741. DOI: [10.1002/nla.801](https://doi.org/10.1002/nla.801).



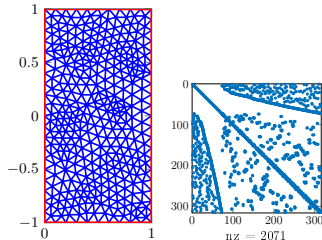
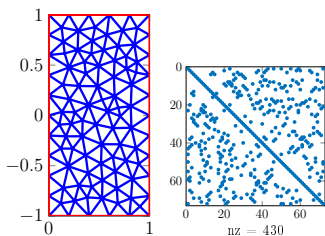
Yuji Nakatsukasa. "Gerschgorin's theorem for generalized eigenvalue problems in the Euclidean metric". *Math. Comp.* 80.276 (2011), S. 2127–2142. DOI: [10.1090/S0025-5718-2011-02482-8](https://doi.org/10.1090/S0025-5718-2011-02482-8).

[EH06] [BR09] [AC85] [BGH13] [CS80] [Nak11] [KVC12]

Matlab: pdetool, refinement regular

Level	n	nnz	nnz/n^2	λ_{\min}	λ_{\max}
1	72	430	0.08295	12.6174	1549
2	315	2071	0.02087	12.4048	8227
3	1317	8965	0.005169	12.3550	3.75e+04
4	5385	3.72e+04	0.001283	12.3415	1.542e+05
5	2.178e+04	1.515e+05	0.0003194	12.3381	6.225e+05
6	8.758e+04	6.112e+05	7.967e-05	12.3373	2.533e+06
7	3.513e+05	2.455e+06	1.989e-05	12.3371	1.023e+07
8	1.407e+06	9.842e+06	4.971e-06	12.3370	4.101e+07

'size', 'nonzeros', 'density', 'real', 'memory', 'sym', 'bandwidth', 'spec'



- Verallgemeinertes Eigenwertproblem, Verallgemeinerung des Satzes von Gershgorin, M strikt diagonaldominant?!
- Kosten Matrix-Vektor-Multiplikation vs. Kosten Lösung des LGS
- Gestalt von M
- Symmetrie, spektrale Eigenschaften von $M^{-1}S$,
- $S = S^H$ spd, Cholesky-Zerlegung / pcg genauso wie für M
- Anzahl m der benötigten Iterationen: Speicher- u. Rechenaufwand wachsen
 - Orthogonalisierung: $\mathcal{O}(m^2) \cdot n$
 - Berechnung $y_m = \varphi_\ell(-S_m)\mathbf{e}_1$ $\mathcal{O}(m^3)$
 - $m + 2$ Vektoren der Länge n
 - Overhead LU -Faktoren \rightarrow fill in trotz Umsortierung
 - LU $\frac{1}{3}n^3 + \mathcal{O}(n^2)$, $\frac{n^2}{2}$ je für vorwärts und rückwärtseinsetzen.
 - Cholesky $\frac{1}{6}n^3 + \mathcal{O}(n^2)$, je für vorwärts und rückwärtseinsetzen.
 - LU - Bandmatrix: npq , p , q obere und untere Bandbreite