Eigenvalue–Eigenvector Structure and Approximation Problems in GMRES Residual Bounds

Jörg Liesen,

CSAR UI, Urbana and TU Berlin,

and

Zdeněk Strakoš,

ICS AS Prague, TU Liberec
and Emory University, Atlanta

XV. Householder Symposium
Peebles, June 2002.

A note for the reader:

A part of the transparencies were skipped during the presentation. The posted text incorporates several changes motivated by the very fruitful discussions with several participants in Peebles.

Maryland, Spring 2000:

Howard Elman, Alison Ramage [FRSW-99]:

GMRES behavior on two-dimensional advection-diffusion problem

$$-\nu \nabla^2 u + w \cdot \nabla u = f$$

in a domain Ω, where the transported quantity u should satisfy some boundary conditions $u = g$ on $\partial\Omega$. For $\Omega = <0,1> \times <0,1>$ and the special case $w = [0,1]$ the SUPG discretization yields the 'SUPG matrices'

$$A = A(\nu, h, \delta) = \text{tridiag}(M_3, M_1, M_2) \in \mathbf{R}^{N^2 \times N^2}.$$

Here $h = 1/(N+1)$ is the discretization parameter, ν the diffusion parameter and δ the 'tuning parameter' which is allowed to vary.

With the SUPG matrices GMRES typically exhibits very slow convergence in the first $N-1$ steps, followed by a significant decrease of the residual norm in the subsequent steps.

Explanation?

Handwaving response:

It is somehow related to the
pattern of cancellation of the possibly large components of the initial residual in the direction of the individual eigenvectors of A.

Emory, Spring 2000:

Jörg Liesen: We should study the linear systems with the SUPG matrices!

Peebles, Wednesday morning, June 19, 2002:

You are sentenced for listening to my report on where we got to.

Outline

1. Initial (transient) phase is important!
2. Eigenvalues and GMRES convergence
3. Bound by Elman step by step
4. SUPG matrices: Eigenvalue-eigenvector structure
5. FRSW problem: Questions
6. FRSW problem: Analysis
7. Closure

1. Initial phase is important!

$Ax = b$, A square and nonsingular.

Iterative methods:

$x_0, x_1, \ldots, x_n, \ldots \longrightarrow x$.

Convergence:

$x_n \approx x$, size of the error $\|x - x_n\|$.

Classical view to the rate of convergence:

$$\lim_{n\to\infty} \frac{\|x - x_{n+1}\|}{\|x - x_n\|^p} = const < 1 ,$$

Practical monitoring of convergence is (unfortunately) typically based on the residual $r_n = \|b - Ax_n\|$.

x_0 , $r_0 = b - Ax_0$.

Krylov subspace methods:

$\mathcal{K}_n \equiv \mathcal{K}_n(A, r_0) \equiv span\,\{r_0, \cdots, A^{n-1} r_0\}.$

$$\begin{aligned} x_n &\in x_0 + \mathcal{K}_n(A, r_0) \,, \\ x - x_n &= p_n(A)\,(x - x_0) \,, \end{aligned}$$

$$\begin{aligned} r_n &\equiv b - Ax_n = p_n(A)\, r_0 \\ &\in r_0 + A\mathcal{K}_n(A, r_0) \,, \qquad p_n(0) = 1 \,. \end{aligned}$$

GMRES method [SS-86]

$$\|r_n\| = \min_{z \in A\mathcal{K}_n(A,r_0)} \|r_0 - z\|$$

How small r_n can be found in the lin. variety

$$r_0 + A\mathcal{K}_n(A, r_0) \ ?$$

On which properties of A, r_0 it depends?

In practical problems, **preconditioning** is absolutely essential. (Which properties should the preconditioned system $Ax = b$ have in order to get fast convergence?)

Restarting when necessary.

Initial approximation: $\|r_n\|/\|r_0\|$ or $\|r_n\|/\|b\|$?

In practical applications, **preconditioned** Krylov subspace methods search for the approximate solution of the finite dimensional problem in

a small number of steps.

In early iterations convergence can strongly depend on the initial residual.

Role of the eigenvalue - eigenvector structure in relation to the particular initial residual ???

Nick Trefethen [Trefethen-97]:

Any use of eigenvalues to derive physical predictions relies on an implicit transformation to eigenvector coordinates. If the matrix is (even moderately) far from normal, the change to eigenvector coordinates may involve an extreme distortion with a superposition of huge eigencomponents that nearly cancel. The state of the system may be determined by the **pattern of cancellation**, rather than by the size of the individual eigen-components.

Without further transformation the eigenvalue - eigenvector structure can hardly be useful!

We will demonstrate that while analysis of the transient phase in terms of eigenvalues and the size of the eigen-components does no good, the cancellation pattern tells the story.

2. Eigenvalues and GMRES

Consider any nonzero eigenvalues
and a sequence (desired convergence curve)

$$f(0) \geq f(1) \geq \ldots \geq f(n-1) > f(n) = 0\,.$$

Then the size of the component of the initial residual eliminated in the jth GMRES step is

$$\phi_j = (f(j-1)^2 - f(j)^2)^{\frac{1}{2}}\,.$$

Let

$$h = (\phi_1, \ldots, \phi_n)^T\,,$$

be the first column of

$$\Phi = \begin{pmatrix} \boxed{h} & \begin{matrix} \boxed{R_{n-1}} \\ 0 \end{matrix} \end{pmatrix},$$

where R_{n-1} is nonsingular upper triangular.

Theorem [APS-97]

The following two assertions are equivalent:

1° *The spectrum of A is $\{\lambda_1, \ldots, \lambda_n\}$ and GMRES(A, r_0) yields residuals such that*

$$\|r_k\| = f(k)\,, \quad k = 0, 1, \ldots, n\,.$$

2° $A = U\Phi C\Phi^{-1}U^*$ *and* $r_0 = Uh$

where C is the companion matrix corresponding to the spectrum of A and U is unitary.

Theorem gives a complete parametrization of the set of **all pairs** $\{A, r_0\}$ for which GMRES gives the prescribed convergence curve while the matrix A has the prescribed eigenvalues.

Many convergence bounds.

A **is diagonalizable:** [Elman-82]

$$\| r_n \| / \|r_0\| \leq \kappa(Y) \min_{p \in \pi_n} \max_i |p(\lambda_i)|$$

polynomial approximation problem
on the spectrum of A, but $\kappa(Y)$?

3. Bound by Elman step by step

A **normal:**

$$
\begin{aligned}
\|r_n\| &= \|p_n(A)r_0\| = \min_{p\in\pi_n} \|p(A)r_0\| \\
&= \min_{p\in\pi_n} \|\, Y \; [p(D)\, Y^* r_0]\| \\
&= \min_{p\in\pi_n} \|p(D)\, Y^* r_0\| \\
&= \min_{p\in\pi_n} \{ \sum_i \mid (y_i^* r_0)\, p(\lambda_i) \mid^2 \}^{\frac{1}{2}} \\
&\le \|r_0\| \; \min_{p\in\pi_n} \max_i |p(\lambda_i)| \, .
\end{aligned}
$$

$p_n(\lambda_i)$ represents a multiplicative correction to the values of the individual components of r_0 in the orthonormal basis $\{y_1, \ldots, y_N\}$ in order to minimize the sum of squares.

A **diagonalizable:**

$$
\begin{aligned}
\|r_n\| &= \|p_n(A)r_0\| = \min_{p\in\pi_n} \|p(A)r_0\| \\
&= \min_{p\in\pi_n} \| Y \, [\, p(D)\, Y^{-1} r_0]\|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \|Y\| \ \min_{p\in\pi_n} \|p(D)\, Y^{-1} r_0\| \\
&= \|Y\| \ \min_{p\in\pi_n} \{\sum_i \mid [Y^{-1} r_0]_i \, p(\lambda_i) \mid^2 \}^{\frac{1}{2}}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \|Y\| \, \|Y^{-1} r_0\| \ \min_{p\in\pi_n} \max_i |p(\lambda_i)| \\
&\leq \|r_0\| \, \kappa(Y) \ \min_{p\in\pi_n} \max_i |p(\lambda_i)| \, .
\end{aligned}
$$

For a general Y, some of the components

$$Y^{-1}r_0$$

can become very large. In such case

$$Y\ [\,p(D)\,Y^{-1}r_0]$$

represents a significant cancellation.
The minimization problem

$$\|r_n\| = \min_{p\in\pi_n} \|\,Y\ [\,p(D)\,Y^{-1}r_0]\|$$

reflects that, while

$$\|Y\|\ \min_{p\in\pi_n} \|\,p(D)\,Y^{-1}r_0\,\|$$

does not (cf. [Trefethen-97]).

A **defective:** $A = SJS^{-1}$,

$$\|r_n\| \ = \ \min_{p \in \pi_n} \| \, S \, [\, p(J) \, S^{-1} r_0] \|$$

The identity and bounds will contain derivatives of p_n at the deffective eigenvalues with degrees determined by the size of the particular Jordan blocks.

Any use?

4. SUPG matrices

$$A = A(\nu, h, \delta) = \text{tridiag}(M_3, M_1, M_2) \in \mathbf{R}^{N^2 \times N^2},$$

where M_1, M_2, M_3 are real, symmetric, tridiagonal and Toeplitz with the diagonal elements $m_{1,3,5}$ and the off-diagonal elements $m_{2,4,6}$ respectively. The eigenvalues-eigenvectors are

$$\begin{aligned}
\lambda_j(M_1) &= m_1 + 2\,m_2 \cos(jh\pi), \quad u_j, \\
\mu_j(M_2) &= m_3 + 2\,m_4 \cos(jh\pi), \quad u_j, \\
\gamma_j(M_3) &= m_5 + 2\,m_6 \cos(jh\pi), \quad u_j,
\end{aligned}$$

for $j = 1, 2, \ldots, N$, where the eigenvector u_j, $\|u_j\| = 1$, is given by

$$u_j = \sqrt{2h} \begin{bmatrix} \sin(jh\pi) \\ \sin(2jh\pi) \\ \vdots \\ \sin(Njh\pi) \end{bmatrix}.$$

Please notice that

$$U \equiv [u_1, \ldots, u_N]$$

is orthonormal and symmetric.

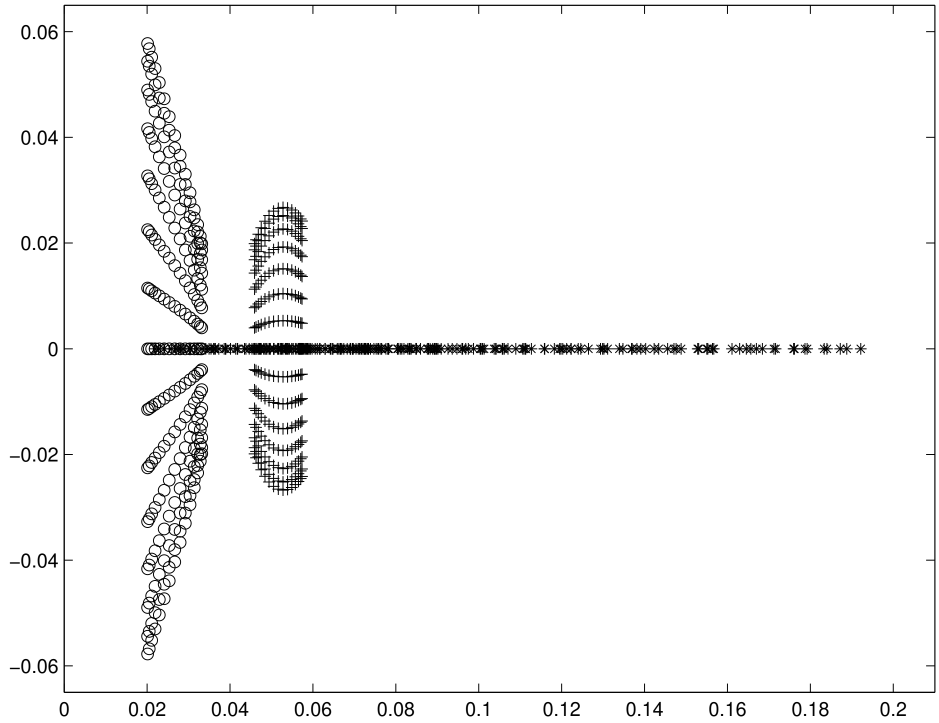
Assuming that M_2 is nonsingular ($\mu_j \neq 0$), the eigendecomposition $A = YDY^{-1}$ is given by

$$\begin{aligned}
A\, y_{jk} &= \sigma_{jk}\, y_{jk}, \quad j,k = 1, 2, \ldots N, \\
\sigma_{jk} &= \lambda_j + 2\sqrt{\gamma_j \mu_j}\, \cos(kh\pi), \\
y_{jk} &= \chi_{jk} \begin{bmatrix} \eta_j \sin(kh\pi)\, u_j \\ \eta_j^2 \sin(2kh\pi)\, u_j \\ \vdots \\ \eta_j^N \sin(Nkh\pi)\, u_j \end{bmatrix}, \;\; \eta_j \equiv \sqrt{\frac{\gamma_j}{\mu_j}},
\end{aligned}$$

where χ_{jk} is chosen such that $\|y_{jk}\| = 1$.

Distribution of the eigenvalues σ_{jk}

(we use $h = 1/16$ $(N = 15)$, $\nu = 1/100$
and $0 \le \delta \le 1$): $\lambda_j > 0$ and $\gamma_j < 0$ for all j.



Eigenvalues of $A(1/100, 1/16, \delta)$
for $\delta = 0$ ('o'), 0.3 ('$+$'), and 0.7 ('$*$').

More on the eigenvalue-eigenvector structure:

Group partitioning

$\mathcal{G}_j \;\equiv\; \{\sigma_{jk},\; y_{jk},\; k = 1, 2, \ldots, N\}, \quad j = 1, \ldots, N.$

Proposition

The eigenvectors in each group $\mathcal{G}_j$ are orthogonal to the ones of each other group.

Inverse eigenvector matrix

Considering $A^T Y^{-T} = Y^{-T} D$ and $A^T =$ tridiag(M_2, M_1, M_3) , we have for $W \equiv Y^{-T}$

$$w_{jk} \;=\; \widehat{\chi}_{jk} \begin{bmatrix} \eta_j^{-1} \sin(kh\pi)\, u_j \\ \eta_j^{-2} \sin(2kh\pi)\, u_j \\ \vdots \\ \eta_j^{-N} \sin(Nkh\pi)\, u_j \end{bmatrix}, \;\; \eta_j \equiv \sqrt{\frac{\gamma_j}{\mu_j}}\,.$$

With this normalization,

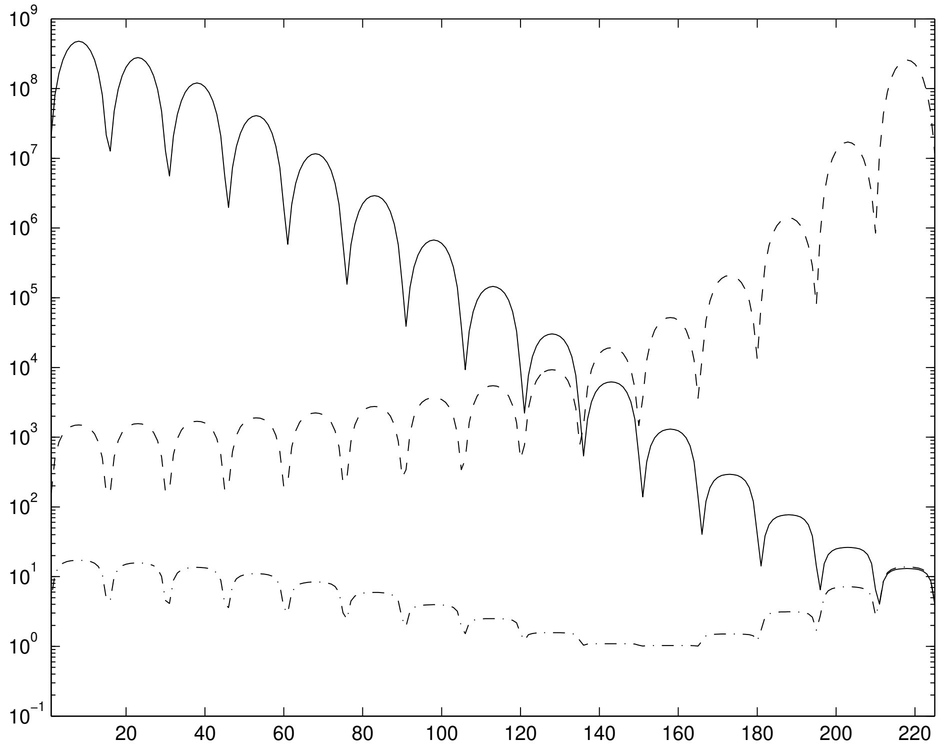
$$\text{cond}(\sigma_{jk}) \;=\; \frac{\|w_{jk}\|\;\|y_{jk}\|}{|w_{jk}^T y_{jk}|} \;=\; \|w_{jk}\|.$$

In our numerical experiments we will use

$$\nu = 1/100, \quad h = 1/16 \;\; (N = 15),$$
$$\text{and} \quad \delta = 0,\; 0.3,\; 0.7\,,$$

and always

$$x_0 = 0 \; (r_0 = b)\,.$$



cond(σ_{jk}) for $A(1/100, 1/16, \delta)$ and $\delta = 0$ (dash-dot), 0.3 (solid), and 0.7 (dashed).

Condition numbers $\kappa(Y)$ (computed with MATLAB's `cond` function)

δ	0	0.3	0.7	1
$\kappa(Y)$	102.6	4.4e+09	2.3e+09	1.4e+04

5. Problem

$$b_P \quad : \quad b \;=\; \text{unit norm PDE related},$$

$$b_Y \quad : \quad b \;=\; (\iota)\, \| \sum_{j,k=1}^{N} y_{jk} \|^{-1} \left(\sum_{j,k=1}^{N} y_{jk} \right),$$

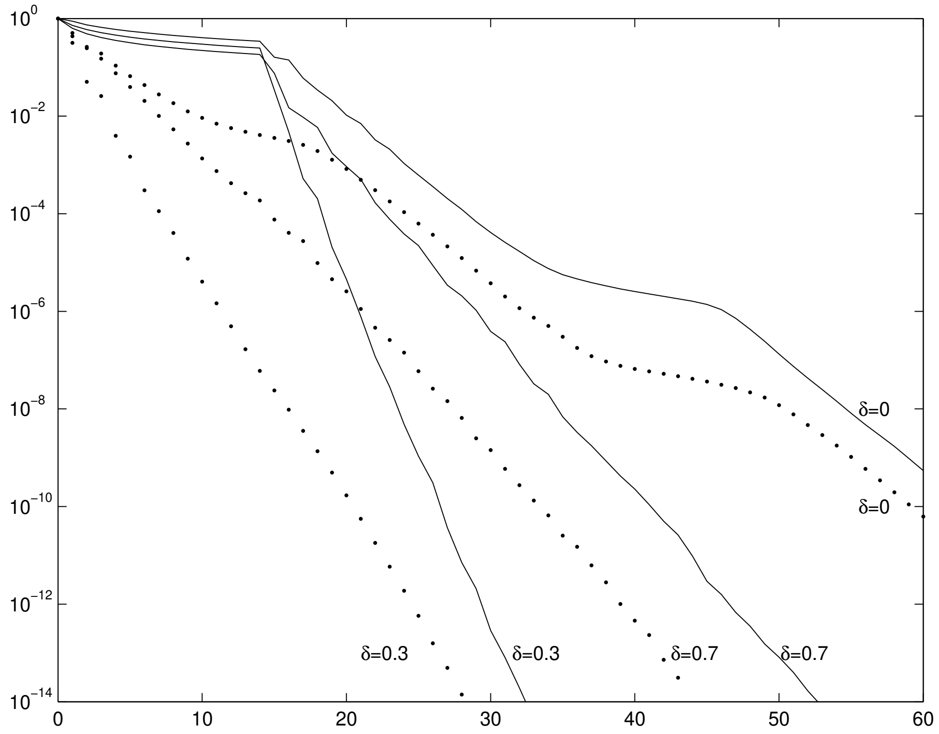
$$b_R \quad : \quad b \;=\; \text{unit norm random},$$

$$b_U \quad : \quad b^{(l)} \;=\; N^{-1} \sum_{j=1}^{N} u_j, \;\; \text{for} \;\; l = 1, \ldots, N,$$

$$b_{U(l)} \quad : \quad b^{(l)} \;=\; N^{-1/2} \sum_{j=1}^{N} u_j, \quad \text{and}$$

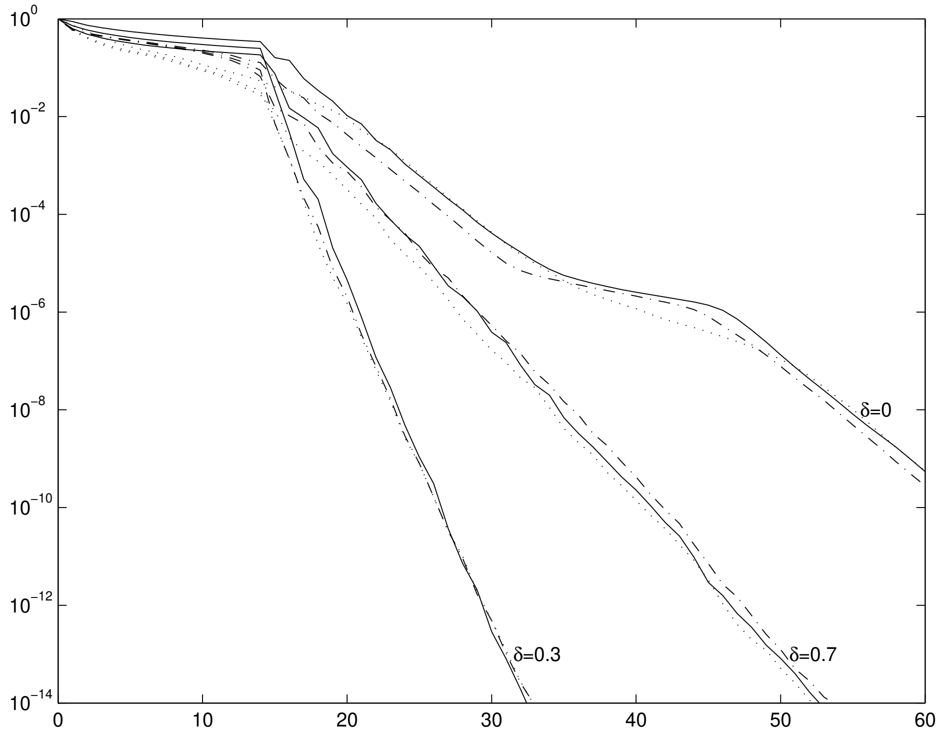
$$b^{(k)} \;=\; 0 \;\; \text{for} \;\; k \neq l.$$

Compare b_P with b_Y :



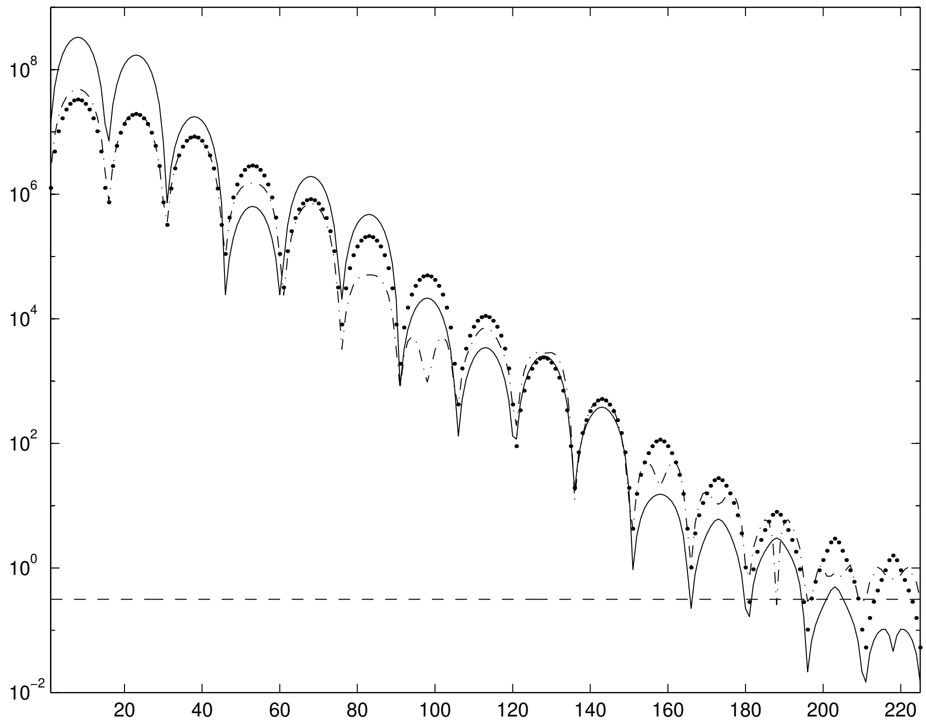
Computed GMRES residual norms $\|r_n\|$ for matrices $A(1/100, 1/16, \delta)$, corresponding right hand sides b_P (solid) and b_Y (dots), $x_0 = 0$, and $\delta = 0$, 0.3, 0.7.

Compare b_P with b_R and b_U :



Computed GMRES residual norms $\|r_n\|$ for matrices $A(1/100, 1/16, \delta)$, corresponding right hand sides b_P (solid), b_R (dotted) and b_U (dashed), and $x_0 = 0$, $\delta = 0$, 0.3, 0.7.

Handwaving explanation: Cancellation of the large components of the initial residual in the direction of the individual eigenvectors



b_P (solid), b_R (dotted) and b_Y (dashed)

6. Analysis

Size of the individual eigen-components is not important! In order to get an insight, we must reveal the cancellation pattern. We will find that in the process the possibly large terms will disappear from our description of convergence.

There is No cancellation between the groups of eigenvectors, and no cancellation between the individual entries of the residual vector. Therefore it is natural to start with two partitionings:

Vertical partitioning

$$r_n = Y\,[\,p_n(D)\,Y^{-1}r_0\,] = \sum_{j=1}^{N} g_{nj},$$

where g_{nj} denotes the contribution of the jth group of eigenvectors

$$g_{nj} = \sum_{k=1}^{N} y_{jk}\,p_n(\sigma_{jk})\,(w_{jk}^T r_0)\,.$$

Due to the mutual orthogonality of the vectors $g_{n1},\ldots,g_{nN}$,

$$\|r_n\|^2 = \sum_{j=1}^{N} \|g_{nj}\|^2\,.$$

Horizontal partitioning

$$r_n \;\equiv\; \begin{bmatrix} r_n^{(1)} \\ r_n^{(2)} \\ \vdots \\ r_n^{(N)} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{N} g_{nj}^{(1)} \\ \sum_{j=1}^{N} g_{nj}^{(2)} \\ \vdots \\ \sum_{j=1}^{N} g_{nj}^{(N)} \end{bmatrix}.$$

and using the structure of y_{jk} and w_{jk},

$$\begin{aligned} g_{nj}^{(m)} &= \sum_{k=1}^{N} y_{jk}^{(m)}\, p_n(\sigma_{jk})\,(w_{jk}^T r_0) \\ &= u_j \sum_{l=1}^{N} c_{nj}^{(ml)}\,(u_j^T r_0^{(l)}) \end{aligned}$$

where

$$c_{nj}^{(ml)} \;\equiv\; \eta_j^{m-l}\, u_m^T\, p_n(D_j)\, u_l$$

and $D_j \equiv \mathrm{diag}(\sigma_{j1}, \ldots, \sigma_{jN})$.

Using the orthonormality of $u_1, \ldots, u_N$

$$\|r_n\|^2 = \sum_{m=1}^{N} \sum_{j=1}^{N} \left| \sum_{l=1}^{N} c_{nj}^{(ml)} \, (u_j^T r_0^{(l)}) \right|^2 .$$

j denotes groups (vertical partitioning),
m denotes blocks (horizontal partitioning).

Initial residual

$$\|r_0\|^2 = \sum_{m=1}^{N} \sum_{j=1}^{N} (u_j^T r_0^{(m)})^2 .$$

Rise of the correction factors

$$c_{nj}^{(ml)} \equiv \eta_j^{m-l} \, u_m^T \, p_n(D_j) \, u_l .$$

In explanation of the main point we use a particular initial residual $(b_{U(l)})$. For a fixed l

$$r_0^{(l)} = N^{-\frac{1}{2}} \sum_{j=1}^{N} u_j\,, \quad \text{and} \quad r_0^{(m)} = 0, \;\; m \neq l\,.$$

The effect of $p_n(D_j)$ on $c_{nj}^{(ml)}$ will be explained for $n = 1$; $p_1(\lambda) = 1 + \rho_1^{(1)}\lambda$.
First

$$D_j u_l \;=\; \lambda_j u_l + \sqrt{\gamma_j \mu_j}\,(u_{l+1} + u_{l-1})\,,$$

where $u_0 = u_{n+1} = 0$.
Starting for each j from a single nonzero correction factor $c_{0j}^{(ll)} = 1$, we get at the first GMRES step a (generally) nonzero correction factors for
$m = l$, $m = l+1$ and $m = l-1$,

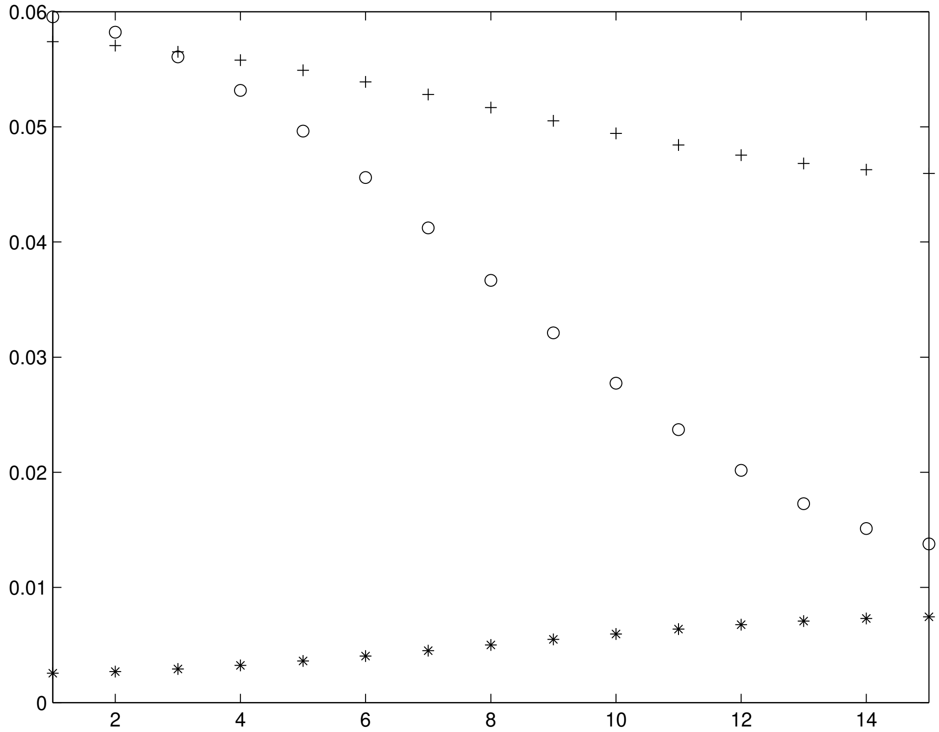
$$
\begin{aligned}
c_{1j}^{(ll)} &= 1+\rho_1^{(1)}\lambda_j\,,\\
c_{1j}^{(l+1l)} &= \rho_1^{(1)}\gamma_j\\
c_{1j}^{(l-1l)} &= \rho_1^{(1)}\mu_j\,.
\end{aligned}
$$

This results (in general) in the three nonzero blocks

$$
\begin{aligned}
\|r_1^{(l)}\|^2 &= \sum_{j=1}^{N} |1+\rho_1^{(1)}\lambda_j|^2\,,\\
\|r_1^{(l+1)}\|^2 &= \sum_{j=1}^{N} |\rho_1^{(1)}\gamma_j|^2\,,\\
\|r_1^{(l-1)}\|^2 &= \sum_{j=1}^{N} |\rho_1^{(1)}\mu_j|^2\,,
\end{aligned}
$$

and all other blocks $r_1^{(m)}$ ($m \neq l$, $l+1$ and $l-1$) of r_1 equal to zero vectors.

The significance (size) of the individual terms is determined by the actual values of λ_j, γ_j and μ_j for $j = 1, \ldots, N$. We consider the particular choice $\nu = 0.01$ and $\delta = 0.3$.



λ_j ('+'), μ_j, ('$*$'), and $|\gamma_j|$ ('o'), $j = 1, \ldots, 15$, for $\nu = 1/100$, $h = 1/16$, and $\delta = 0.3$.

Unless $l = N$, any possible significant drop of $\|r_1^{(l)}\|$ *is compensated* by a concurrent significant rise of $\|r_1^{(l+1)}\|$, which prevents a significant decrease of $\|r_1\|$ in comparison to $\|r_0\|$. Consequently, unless $l = N$, the convergence of GMRES in the first step must be slow. Please notice that for the given ν and δ the effect of $\|r_1^{(l-1)}\|$ is totally insignificant, because $\mu_j \ll |\gamma_j|$.

Summarizing, with $\nu = 0.01$ and $\delta = 0.3$ the rise of the new nonzero block $r_1^{(l+1)}$ in the forward direction $(l \to l + 1)$ is responsible for preventing fast convergence in the first step, while the rise of the new nonzero block $r_1^{(l-1)}$ in the backward direction $(l \to l - 1)$ is insignificant.

Similar considerations apply to a general GMRES iteration step n, $n = 2, \ldots, N$.

Any possible significant drop of $\|r_n^{(l)}\|$ (and of the forward blocks $\|r_n^{(l+1)}\|, \ldots, \|r_n^{(l+n-1)}\|$ modified from the nonzero blocks rised at the preceeding steps) is compensated by a comparable rise of the new forward block

$$\|r_n^{(l+n)}\|^2 = \sum_{j=1}^{n} |\rho_n^{(n)} \gamma_j^n|^2 \, .$$

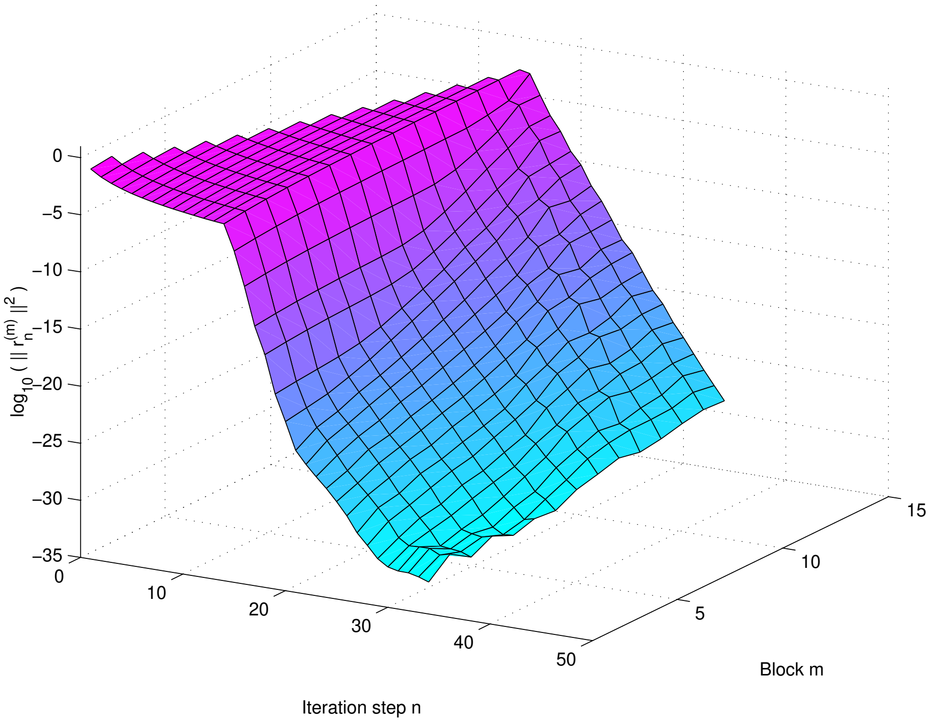
As for $n = 1$, the backward blocks $r_n^{(l-n)}$

$$\|r_n^{(l-n)}\|^2 = \sum_{j=1}^{n} |\rho_n^{(n)} \mu_j^n|^2$$

are, due to the relations between μ_j, λ_j and γ_j (for the chosen ν and δ) insignificant.

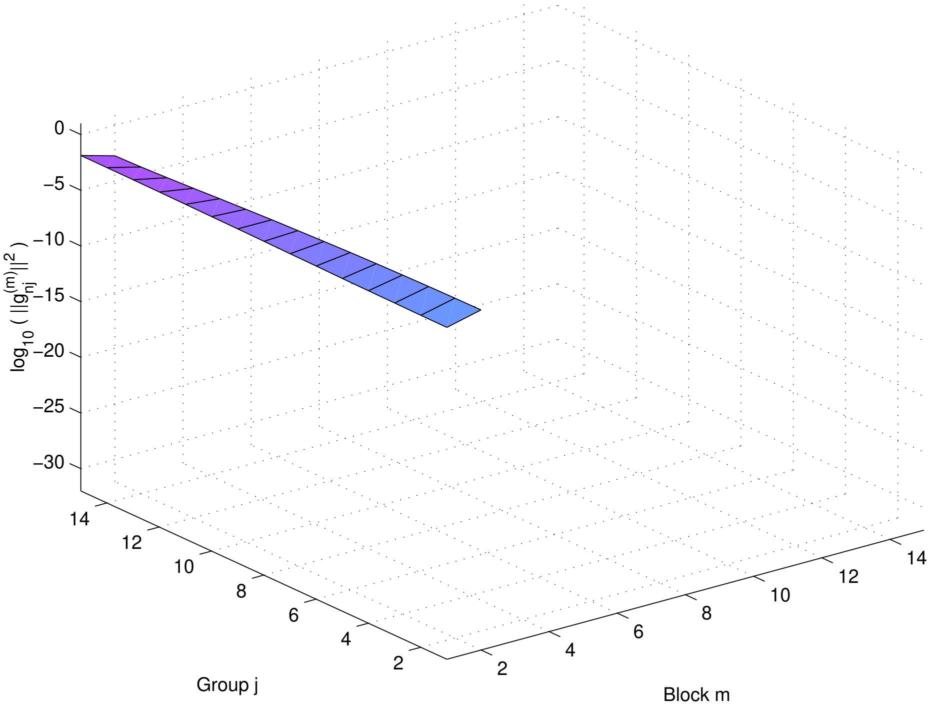
With a given fixed l and $r_0^{(l)}$ the only nonzero block of r_0, we can therefore expect $n-l$ steps of slow GMRES convergence, followed by a considerably increased rate of convergence at the subsequent steps.

$l = 1,\ \delta = 0.3$

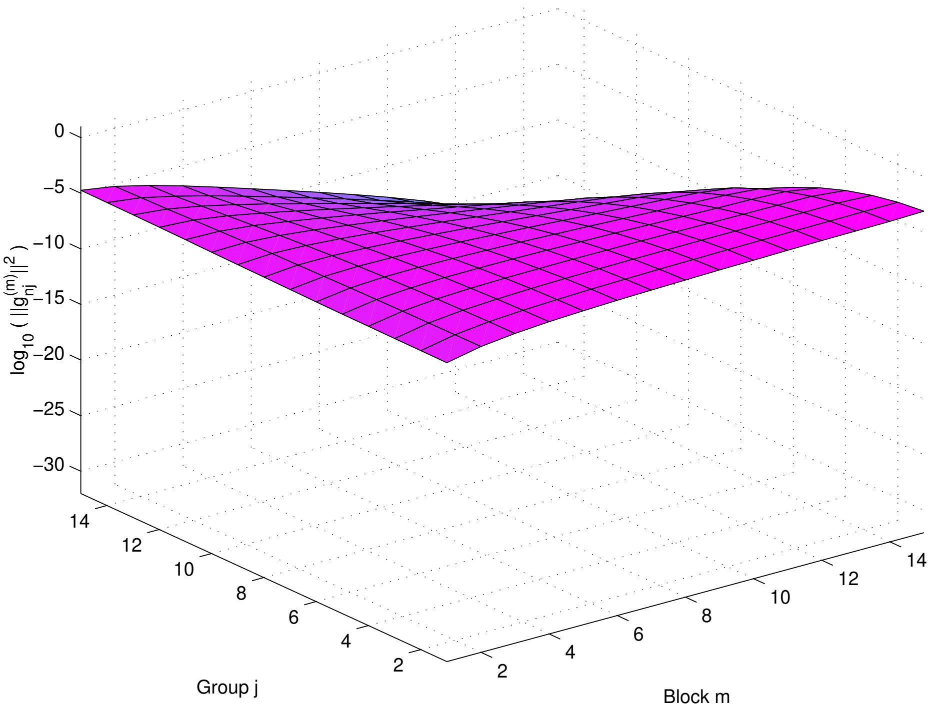


Group partitioning $\|g_{nj}^{(m)}\|^2$ of the residual blocks $\|r_n^{(m)}\|^2$

$$n = 1$$

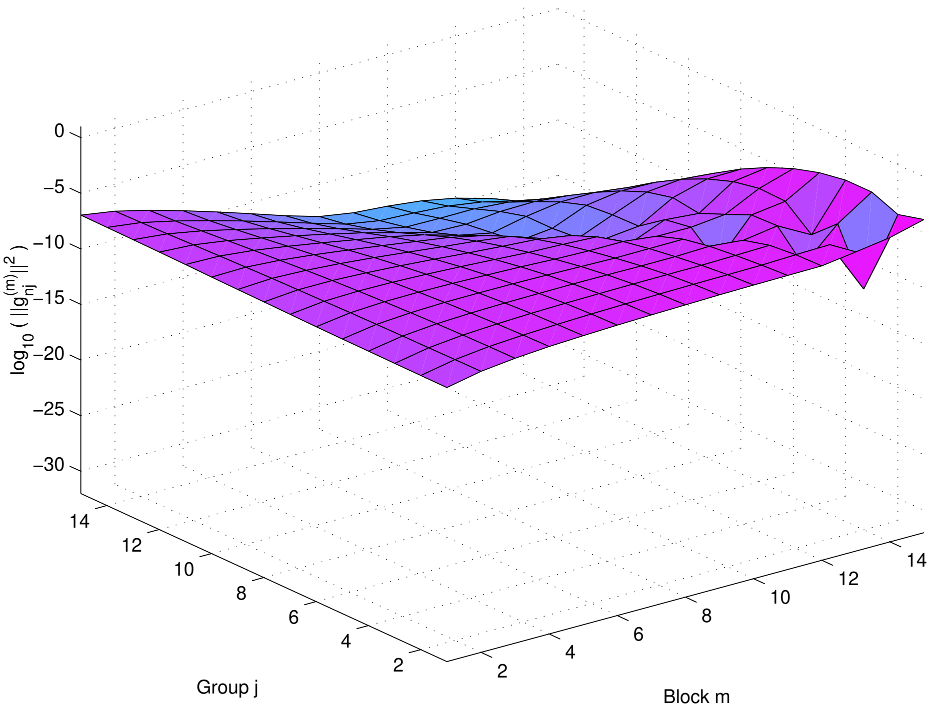


$$n = 14$$



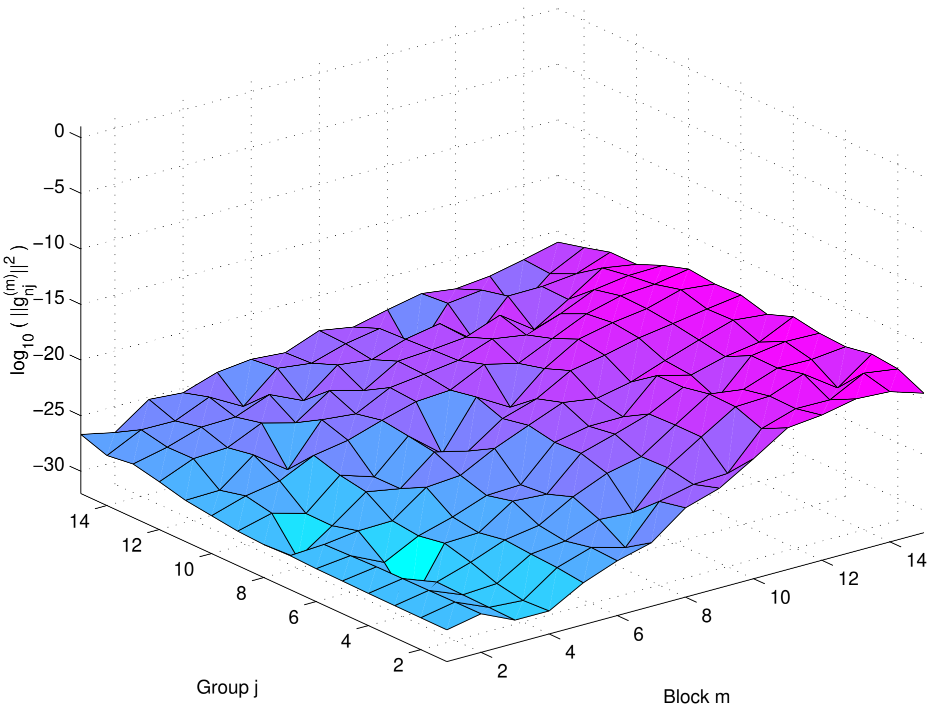
0
−5
−10
−15
−20
−25
−30
log10 (||g(m)nj||2)
14
12
10
8
6
4
2
Group j
2
4
6
8
10
12
14
Block m

$n = 15$



0
−5
−10
−15
−20
−25
−30
$\log_{10}(\,||g_{nj}^{(m)}||^2\,)$
14
12
10
8
6
4
2
Group j
2
4
6
8
10
12
14
Block m

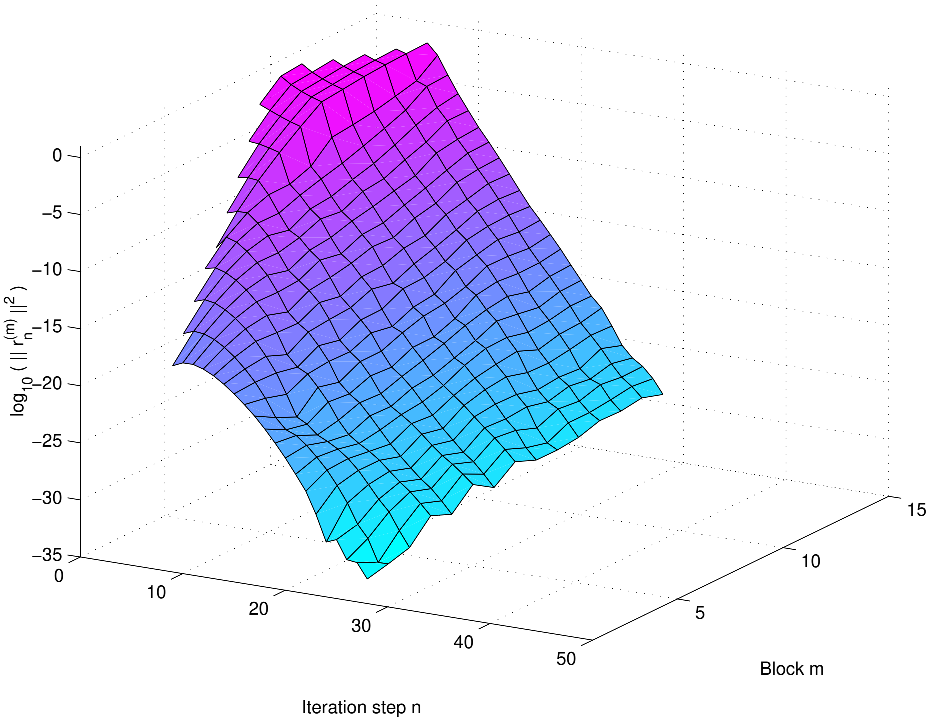
$n = 25$



0
−5
−10
−15
−20
−25
−30
$\log_{10}(\|g_{nj}^{(m)}\|^2)$
14
12
10
8
6
4
2
Group j
2
4
6
8
10
12
14
Block m

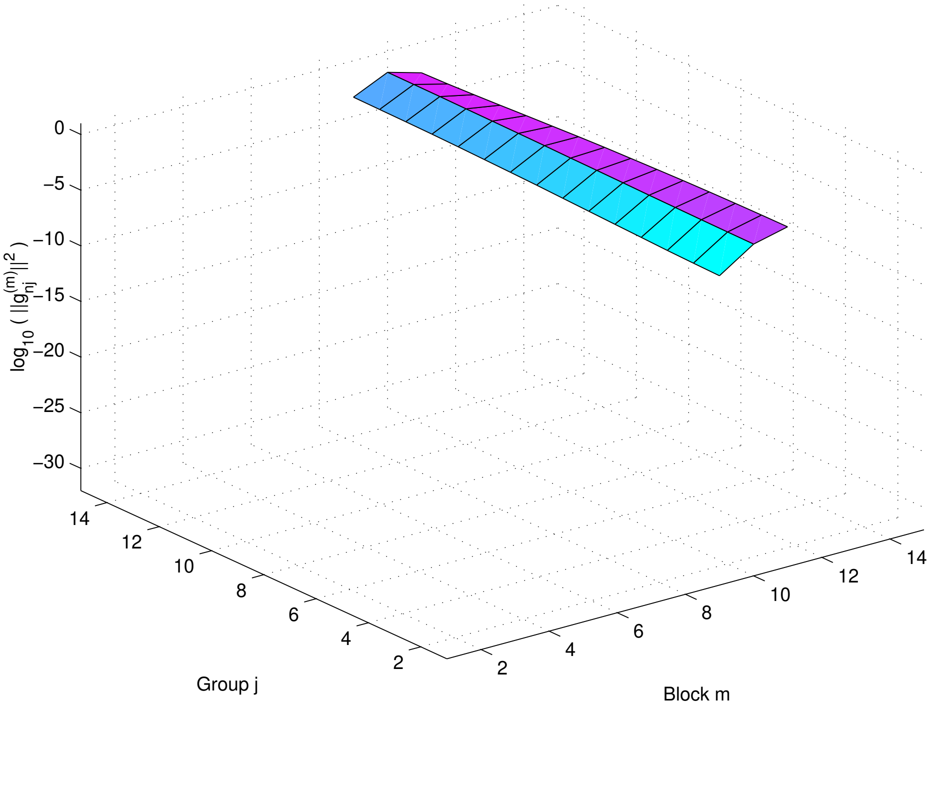
Residual blocks $\|r_n^{(m)}\|^2$

$l = 10,\ \delta = 0.3$

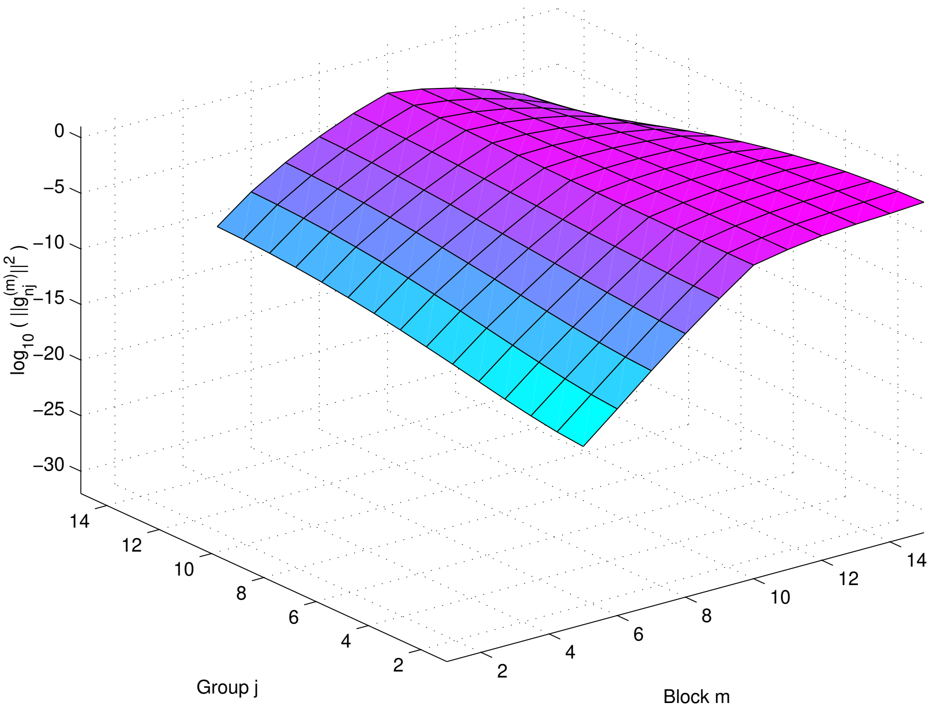


Group partitioning $\|g_{nj}^{(m)}\|^2$ of the residual blocks $\|r_n^{(m)}\|^2$

$n = 1$

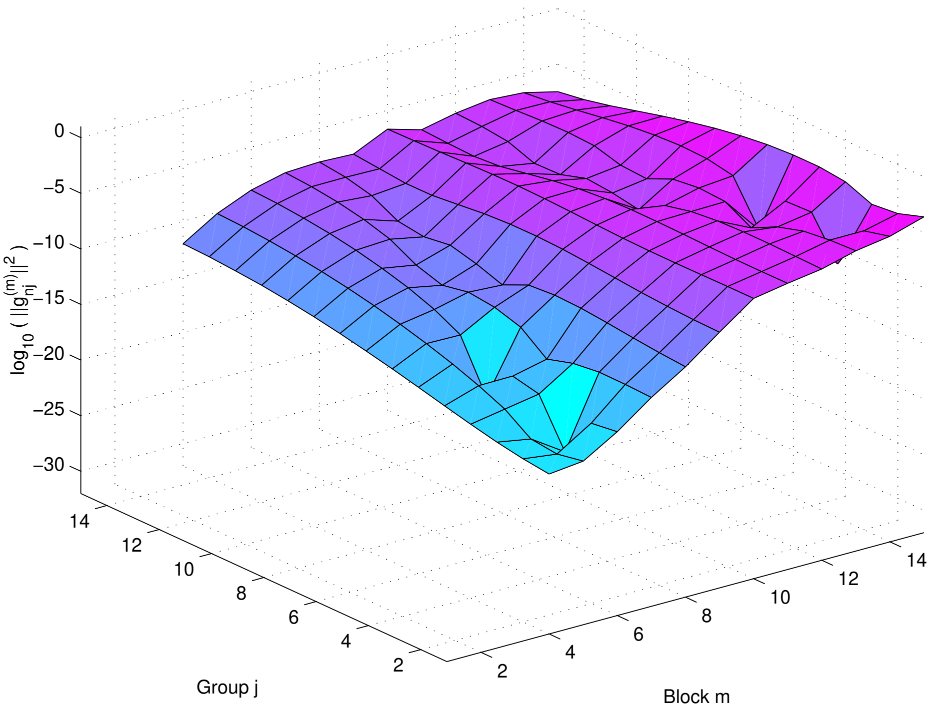


$$n = 5$$



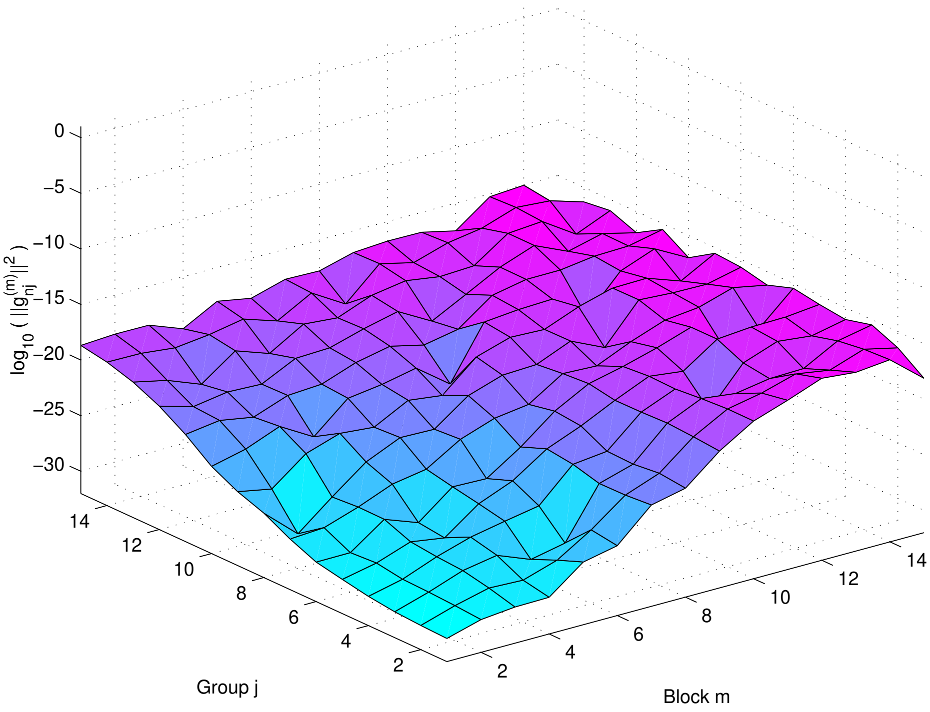
0
−5
−10
−15
−20
−25
−30
$\log_{10}(\,\|g_{nj}^{(m)}\|^2\,)$
14
12
10
8
6
4
2
Group j
2
4
6
8
10
12
14
Block m

$n = 6$



0
−5
−10
−15
−20
−25
−30
$\log_{10}(\|g_{nj}^{(m)}\|^2)$
14
12
10
8
6
4
2
Group j
2
4
6
8
10
12
14
Block m

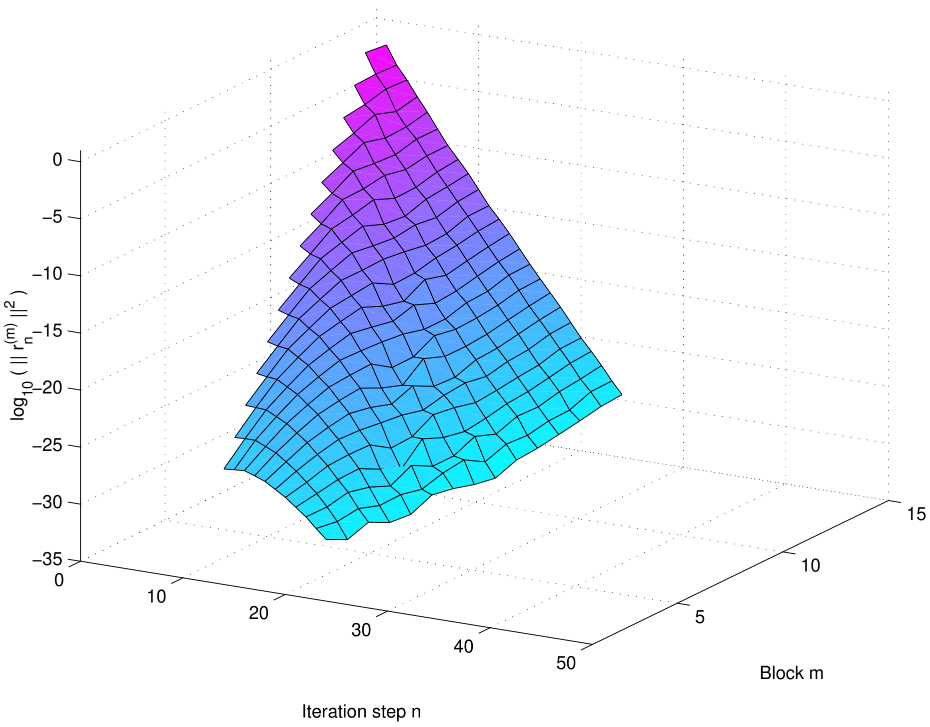
$n = 15$



0
−5
−10
−15
−20
−25
−30
log10 (||g(m)nj||2)
14
12
10
8
6
4
2
Group j
2
4
6
8
10
12
14
Block m

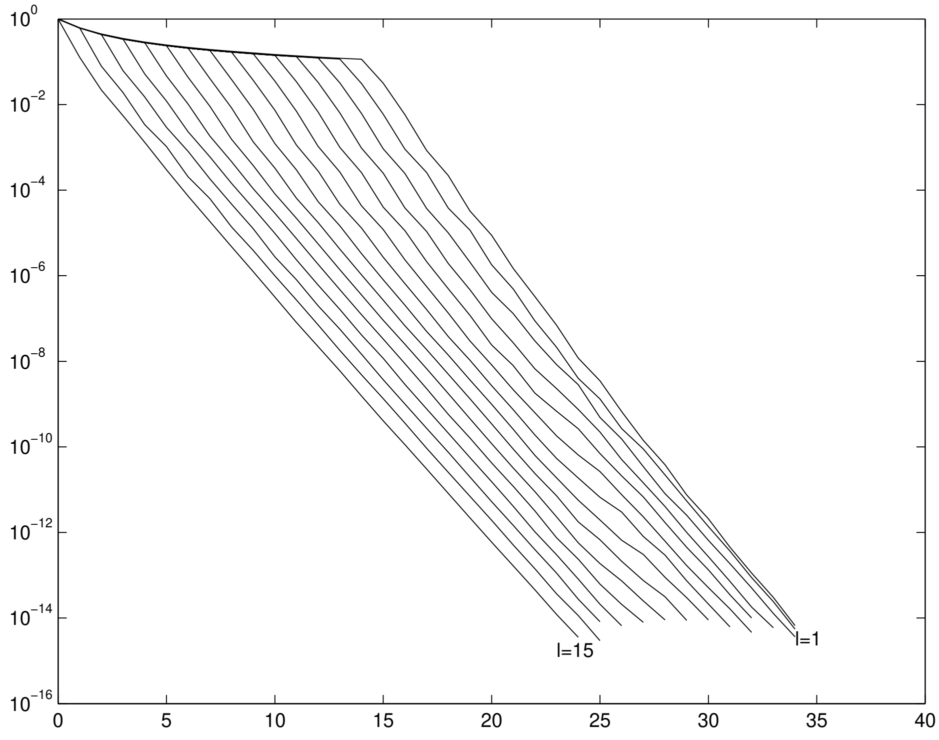
Residual blocks $\|r_n^{(m)}\|^2$

$l = 14,\ \delta = 0.3$

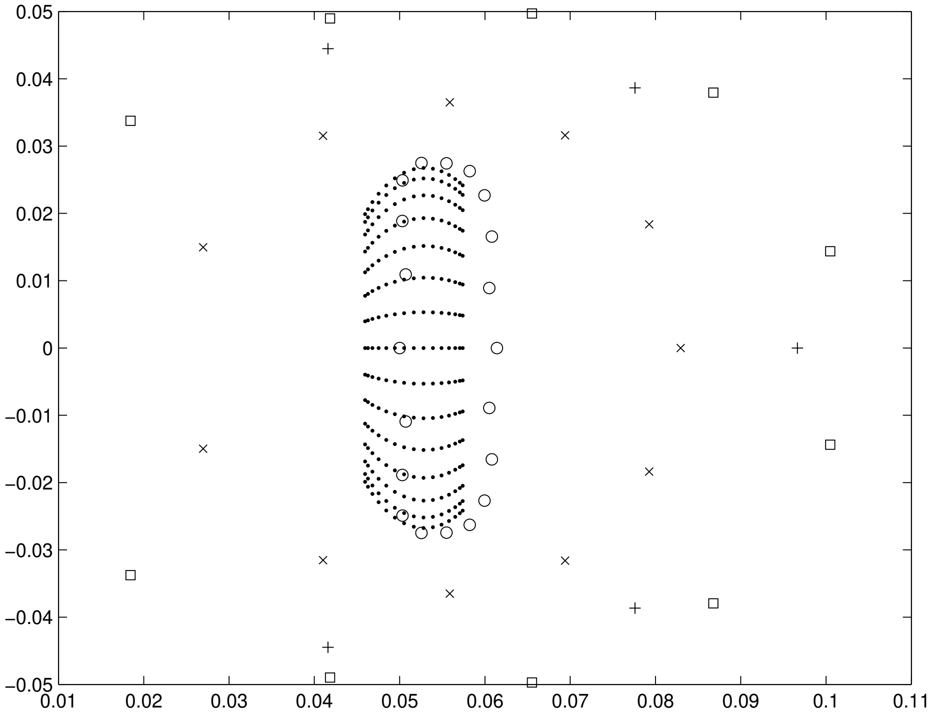


Residual norm for $b_{U(l)}$ with different l

$$l = 1, 2, \ldots, 15, \delta = 0.3$$



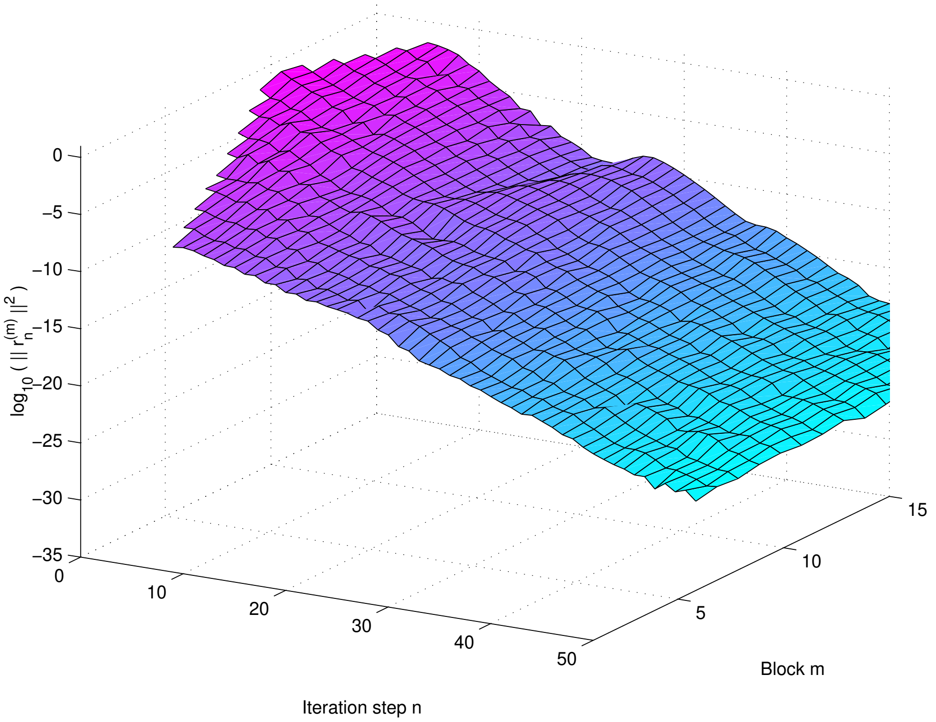
Roots of the GMRES polynomial, $l = 5$

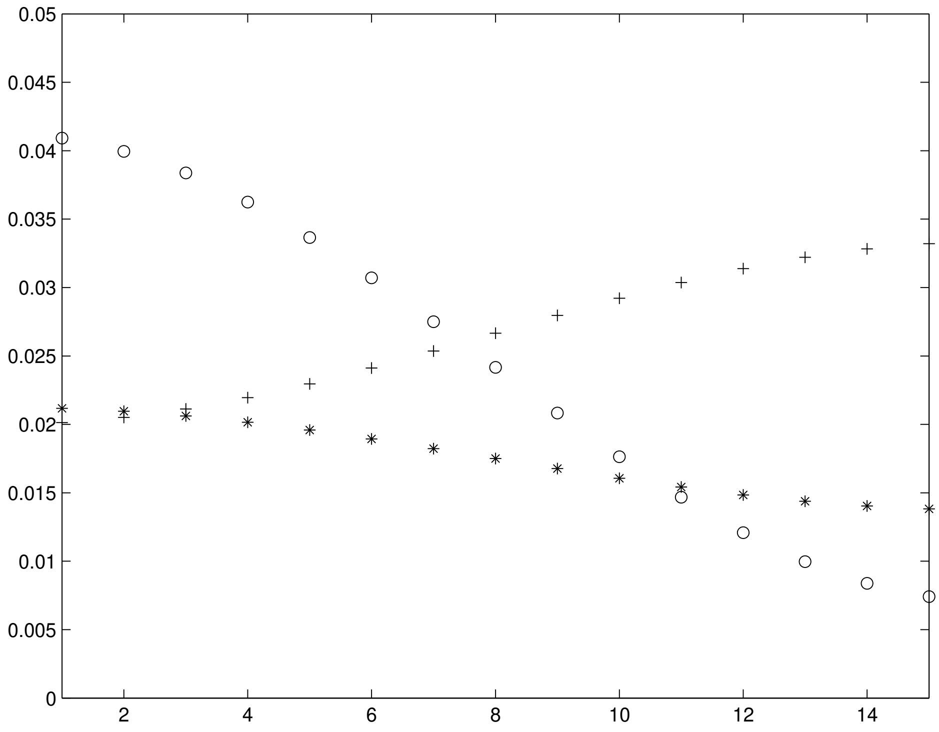


$n = 5(`+`), n = 10(`\square`), n = 11(`\times`)$ and $n = 20(`\circ`)$

Residual blocks $\|r_n^{(m)}\|^2$

$l = 10,\ \delta = 0$

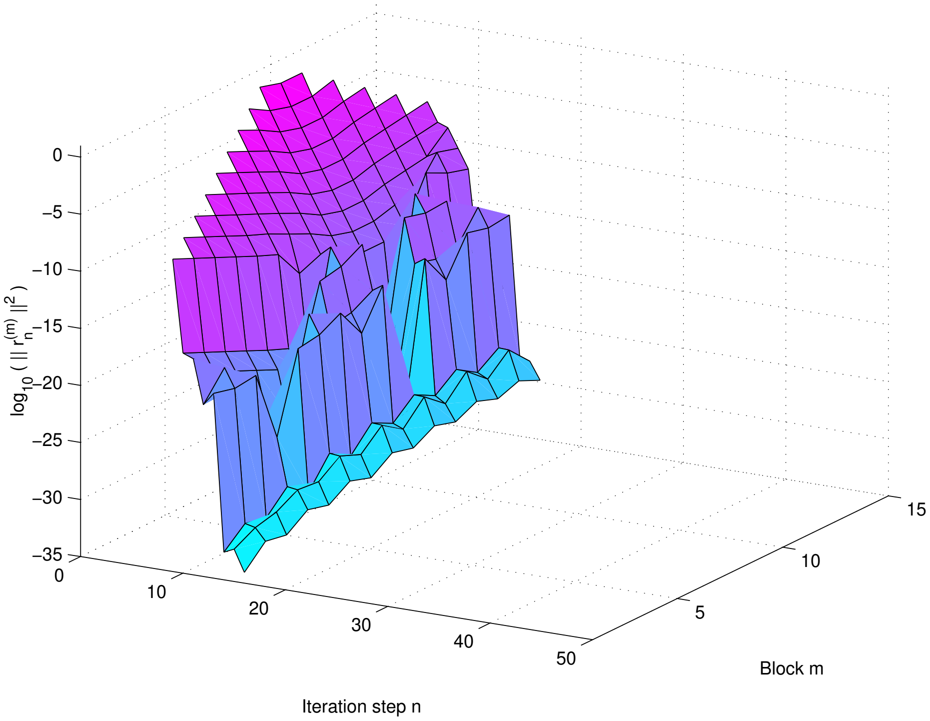




λ_j ('+'), μ_j, ('$*$'), and $|\gamma_j|$ ('o'), $j = 1, \ldots, 15$,
for $\nu = 1/100$, $h = 1/16$, and $\delta = 0$.

Residual blocks $\|r_n^{(m)}\|^2$

$l = 10,\ \delta = 0,\ b^{(10)} = u_{15}$



A generally **nonzero blocks for all l:**

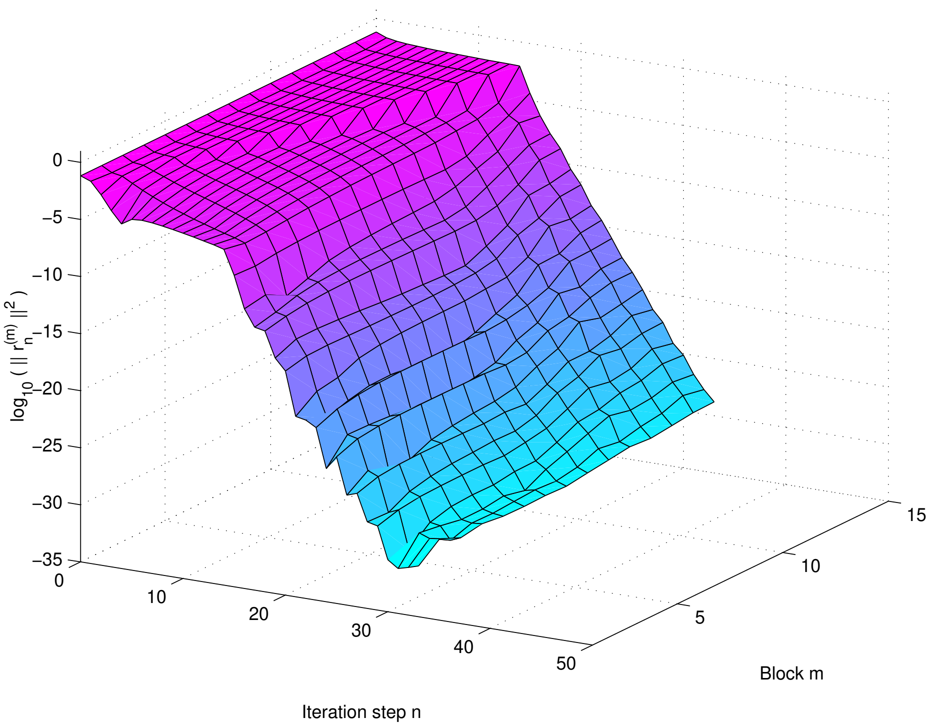
$$\|r_n^{(m)}\|^2 = \sum_{j=1}^{N} \left| \sum_{l=1}^{N} c_{nj}^{(ml)} \, (u_j^T r_0^{(l)}) \right|^2 .$$

Each block starts with a single term in the inner sum, and gets new terms per each iteration until the sum is complete. For $m = 1$ and $m = 15$ it takes 14 iterations.

Forward and backward directions are not of the same importance.

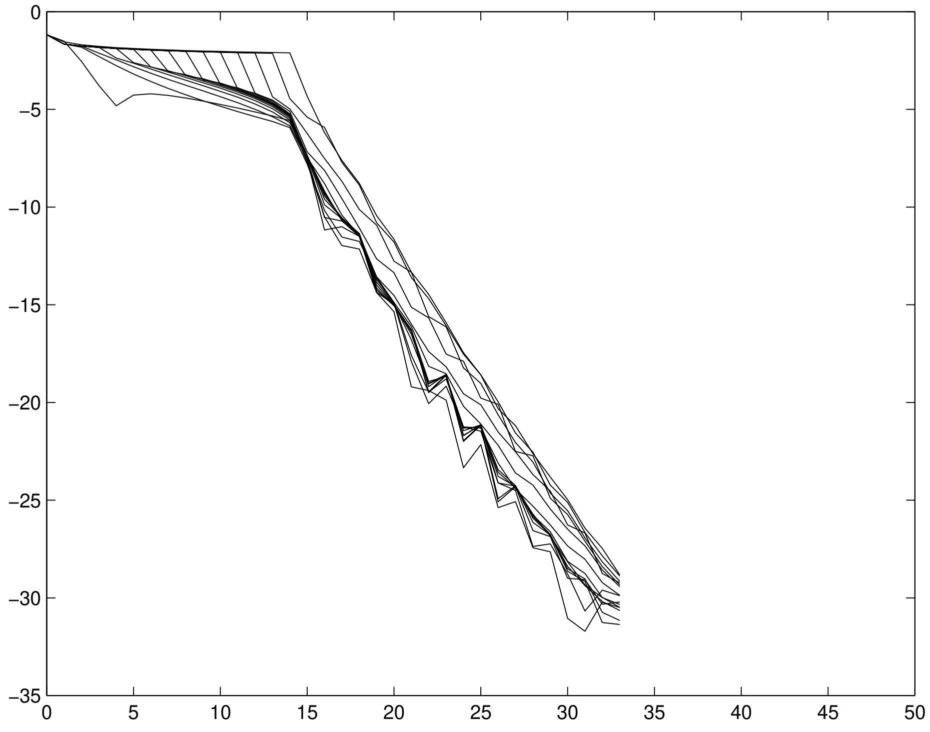
Residual blocks $\|r_n^{(m)}\|^2$

$b_U,\ \delta = 0.3$



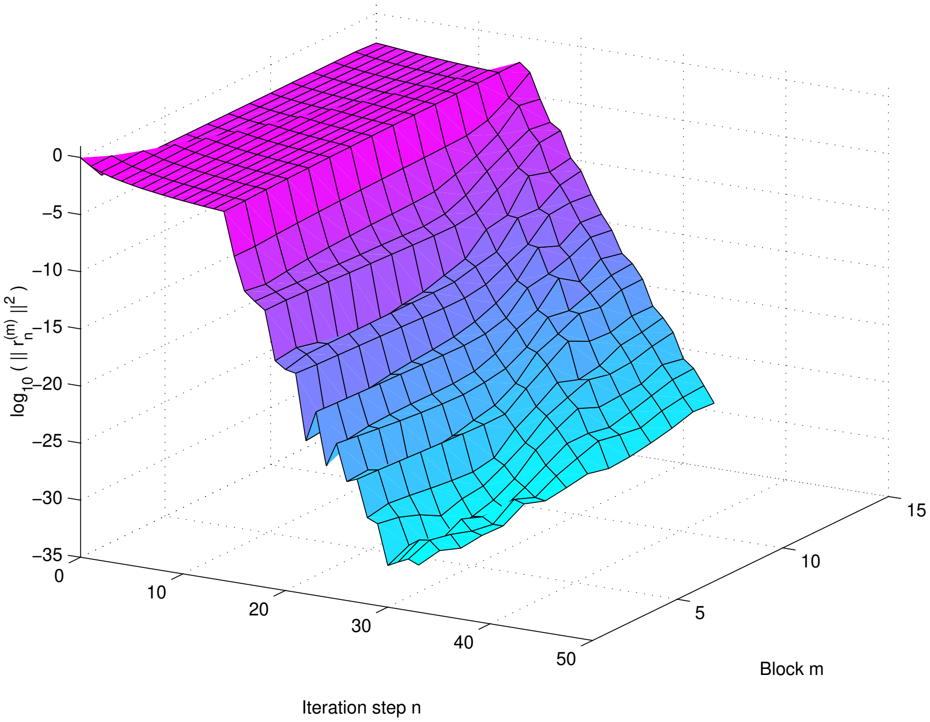
Individual blocks $\|r_n^{(m)}\|^2$

$m = 1, \ldots, 15$



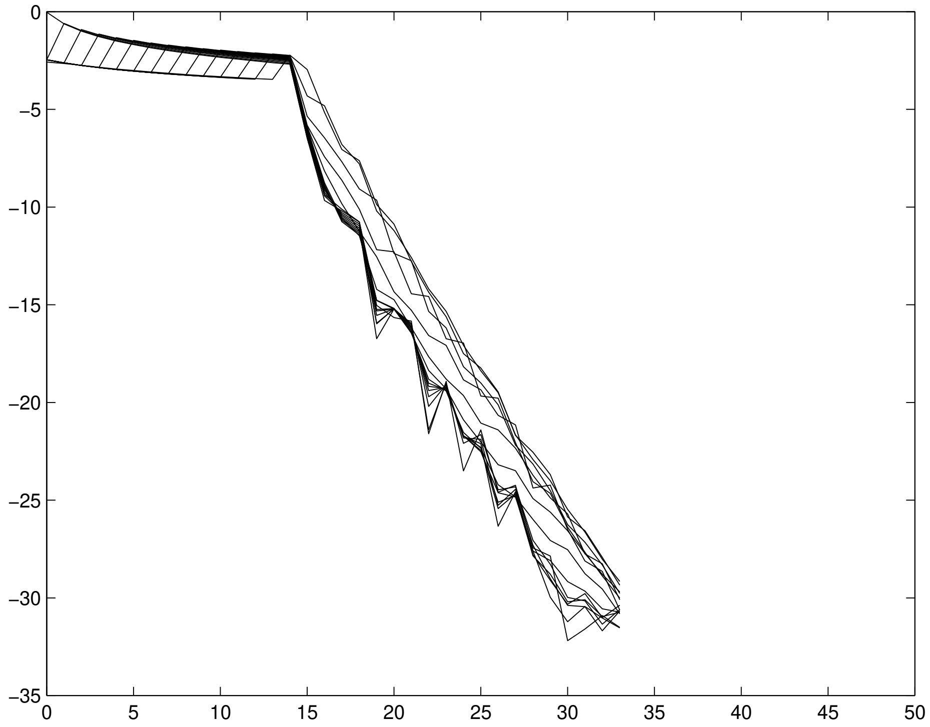
Residual blocks $\|r_n^{(m)}\|^2$

$b_P, \ \delta = 0.3$



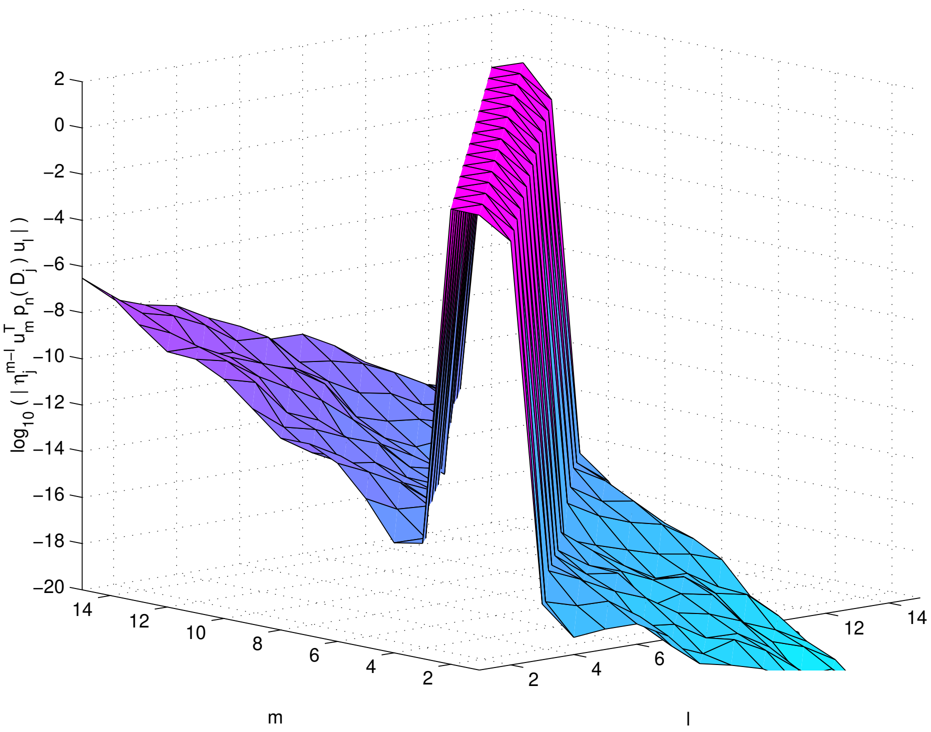
Individual blocks $\|r_n^{(m)}\|^2$

$m = 1, \ldots, 15$

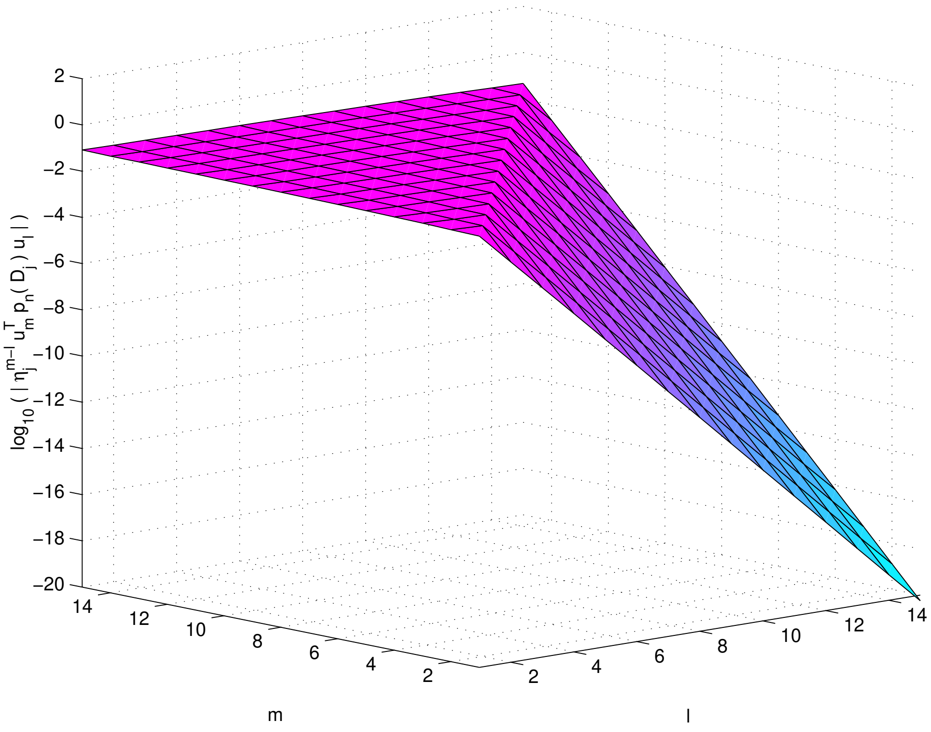


Correction factors for b_P, $\delta = 0.3$, $j = 1$

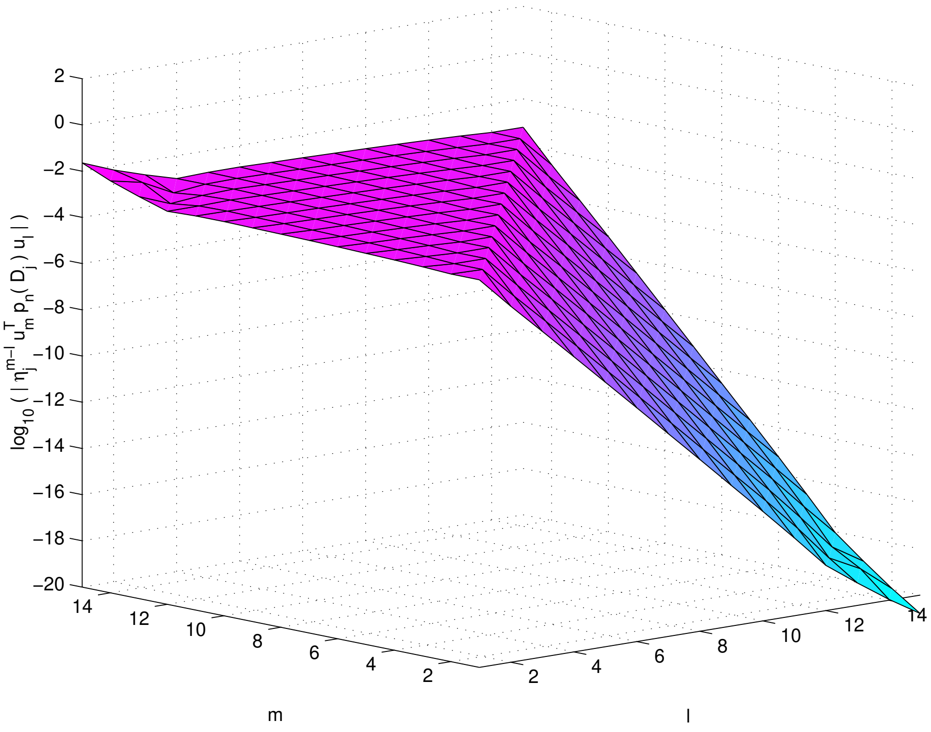
$n = 1$



$n = 14$

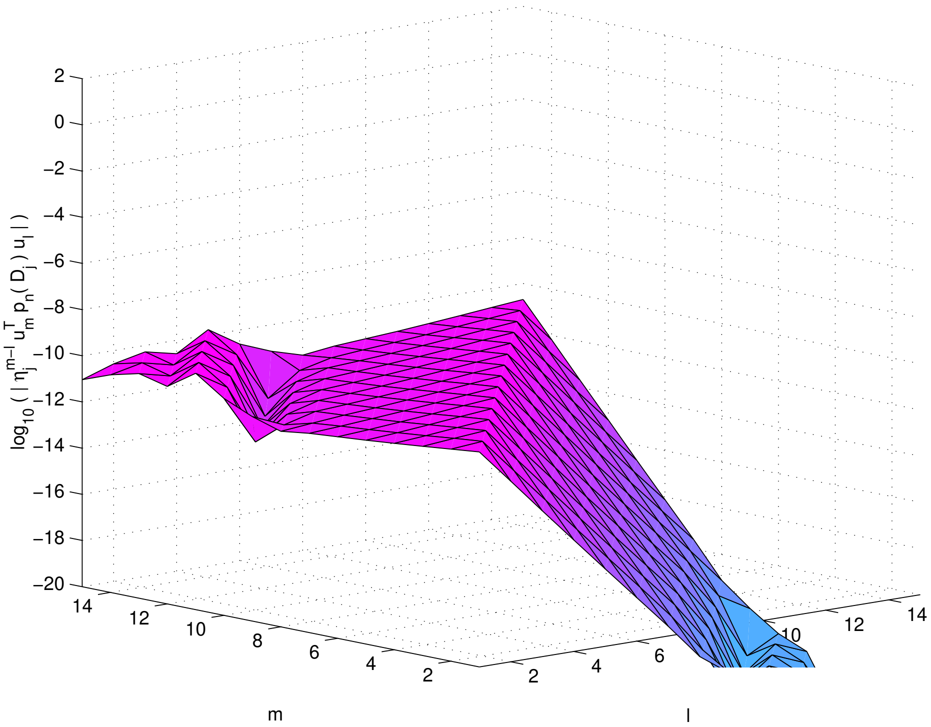


$n = 15$



$\log_{10}(\,|\,\eta_j^{m-l}\,u_m^T p_n(\,D_j\,)\,u_l\,|\,)$
2
0
−2
−4
−6
−8
−10
−12
−14
−16
−18
−20
14
12
10
8
6
4
2
m
2
4
6
8
10
12
14
l

$n = 25$

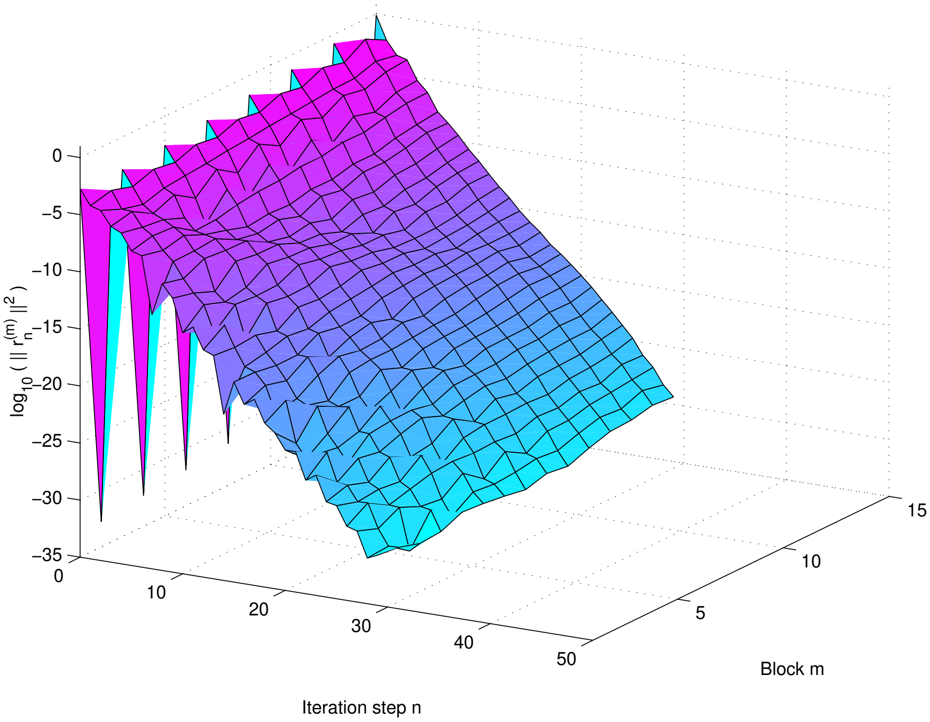


2
0
−2
−4
−6
−8
−10
−12
−14
−16
−18
−20
log10(| ηj m−l um T pn(Dj) ul |)
14
12
10
8
6
4
2
m
2
4
6
10
12
14
l

b_Y $(\delta = 0.3)$ **why no stagnation?**

$$r_0^{(2)} = r_0^{(4)} = \ldots = r_0^{2\lfloor N/2 \rfloor} = 0$$

$$r_0^{(1)} \ll r_0^{(3)} \ll \ldots \ll r_0^{2\lceil N/2 \rceil - 1}$$



7. Closure

- SUPG matrices: when (for which right hand sides), why and for how many steps slow initial convergence.

- Asymptotic convergence: FRSW paper
  $b_P \times b_Y$?

- Physical interpretation?

- Any generalization?

- At least some nice figures!
  Compliments to MATLAB.