Symbol approach (for structured matrices) in multigrid with applications in imaging

Marco Donatelli

Dipartimento di Fisica e Matematica
Università dell'Insubria - Como

Collaborators: *S. Serra-Capizzano, A. Aricò*

"EMG 2010"
Ischia 19–23 September, 2010

Main Issues

Toeplitz matrices and V-cycle optimality

1. Structured matrices and the symbol: spectral analysis
2. Preconditioning and negative results
3. Multigrid methods for Toeplitz matrices
4. V-cycle optimality via a symbol approach

Deblurring of images

5. Deblurring problems
6. Tikhonov regularization
7. Multigrid Regularization

Examples of structures .... and of hidden structures

- Toeplitz, Circulants , τ, Hartley, trigonometric algebras, anti-reflective algebra, algebra associated to wavelet transforms, α-Toeplitz, α-circulants (with blocks and/or with multilevel structure)
- Locally Toeplitz, Generalized Locally Toeplitz (with blocks and/or with multilevel structure)

Toeplitz Sequences: shift-invariance

- $f \in L^1(I_d)$, $I_d = (-\pi, \pi)^d$, $\mathbf{i}^2 = -1$, $j \in \mathbb{Z}^d$;
- $a_j = \frac{1}{(2\pi)^d} \int_{I_d} f(s) e^{-\mathbf{i} js} \, ds$.

For $d = 1$:

$$T_n(f) = \begin{pmatrix} a_0 & a_{-1} & \cdots & a_{1-n} \\ a_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{-1} \\ a_{n-1} & \cdots & a_1 & a_0 \end{pmatrix},$$

The Toeplitz matrix $T_n(f)$ of size N, $n = (n_1, \ldots, n_d)$, $N = \prod n_j$, is defined as

$$T_n(f) = \sum_{|j| \le n-1} a_j J_n^{[j]}, \qquad J_n^{[j]} = J_{n_1}^{j_1} \otimes \cdots \otimes J_{n_d}^{j_d},$$

with $(J_m^r)_{s,t} = 1$ if $s - t = r$ and 0 otherwise.

Circulants: the natural approximation for Toeplitz

$$C_n(f) = \begin{pmatrix} a_0 & a_{n-1} & \cdots & a_1 \\ a_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1} \\ a_{n-1} & \cdots & a_1 & a_0 \end{pmatrix} = F_n^* \mathrm{diag}(\sqrt{n} F_n \mathbf{a}) F_n,$$

$$F_n = \frac{1}{\sqrt{n}} \left[\mathrm{e}^{\frac{-\mathbf{i}2\pi jk}{n}} \right]_{j,k=0}^{n-1}.$$

- $y = C_n(f)\mathbf{x}$ computable within $O(n \log(n))$ operations (also inversion, eigenvalue computation, etc.),
- Multilevel (block) version
- Other Sine, Cosine, ... algebras related to different BCs.

Spectral Distributions

- The eigenvalues of the Circulant matrix $C_n(f)$ are

$$\boldsymbol{\lambda} = \sqrt{n} F_n \mathbf{a} = f(\mathbf{z})$$
$$z_k = \frac{2\pi k}{n}, \qquad k = 0, \dots, n-1.$$

- The eigenvalues of the Toeplitz matrix $T_n(f)$ are "similarly" distributed:

$$\{T_n(f)\} \sim_\lambda (f, (-\pi, \pi)),$$

- The same holds in the d-dimensional case.

Toeplitz Sequences: spectral analysis via the symbol

- $f(s) = (2 - 2\cos(s))^2 = 6 - 4e^{\mathbf{i}s} - 4e^{-\mathbf{i}s} + e^{2\mathbf{i}s} + e^{-2\mathbf{i}s}$ (unique zero at $s = 0$ of order 4)

$$T_n(f) = \begin{pmatrix} 6 & -4 & 1 & & \\ -4 & \ddots & \ddots & \ddots & \\ 1 & \ddots & & \ddots & 1 \\ & \ddots & \ddots & \ddots & -4 \\ & & 1 & -4 & 6 \end{pmatrix},$$

$$\{T_n(f)\} \sim_\lambda (f, (-\pi, \pi)),$$

$$\kappa_2(T_n(f)) \sim n^4.$$

- $f \geq 0$, $\inf f = 0$ (finite number of zeros): $\kappa_2(T_n(f)) \sim N^{\alpha/d}$, α maximal order of the zeros (Szegö, Böttcher, Silbermann, Tyrtyshnikov-Zamarashkin, Tilli, Serra-Capizzano, ...).

GLT Sequences and Approximate PDEs

The structures represent properties inherited from the original physical problem modeled and/or approximated in matrix form:

$$-\nabla^T[a(x)\nabla u]+p(x)\nabla u+q(x)u=g(x), \quad x\in\Omega\subset\mathbb{R}^2, \text{ proper BCs.}$$

$A_n \in \mathbb{R}^{d_n\times d_n}$ = Finite Differences or Finite Elements on a uniform grid discretization or triangulation.

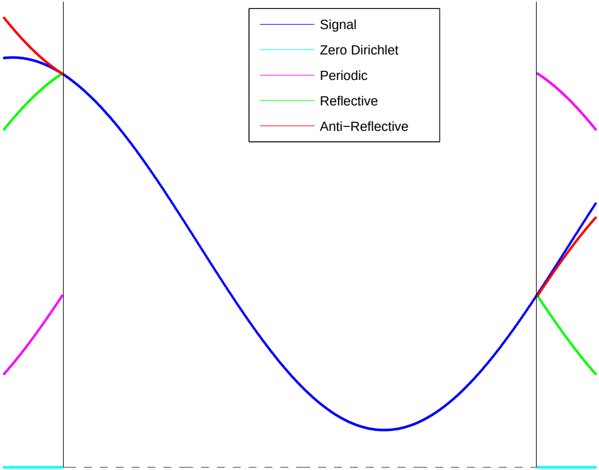
- What about the eigenvalues of A_n? Are they evaluations of a function θ on a uniform gridding of a given domain D?

$$\{A_n\}\sim_\lambda(\theta,D), \quad D=\Omega\times I_2, \quad I_2=(-\pi,\pi)^2$$

$$\theta(x,s)=a(x)(4-2\cos(s_1)-2\cos(s_2)), \quad \kappa_2(A_n)\sim d_n.$$

- We need only Riemann integrability of $a(x)$ and Peano-Jordan measurability of Ω.
- Generalization to general PDEs [Tilli, LAA '98, Serra-Capizzano, LAA '03, ...].

Boundary Conditions for Signals



Various related algebras: Ng, R.Chan, Tang, SISC'99, Serra-Capizzano, SISC'03

Preconditioning

Given the system (to be solved by a Krylov method)

$$A_n x = b,$$

find P_n and solve $P_n^{-1} A_n x = P_n^{-1} b$ such that

Optimality:

- the solution of a system with matrix P_n can be performed within the cost of a matrix-vector product with the matrix A_n,
- $P_n^{-1} A_n$ is a "good spectral approximation" of the identity (constant number of iterations).

Circulant preconditioning of Toeplitz matrices

$d = 1$

Optimal preconditioner for $T_n(f)$ can be defined in the set of Circulant matrices (or other matrix algebras) [T. Chan, SISC '88].

$d > 1$

We should not expect the classical matrix algebra preconditioners to lead to optimal iterative solvers in the ill-conditioned case.

When using matrix Algebra preconditioners, we lose necessarily the optimality and the number of iterations grows like $O(N^{\frac{d-1}{d}})$ [Serra-Capizzano, Tyrtyshnikov, SIMAX '99].

$\Downarrow$

Multigrid methods!

Multigrid methods for Toeplitz matrices

Assumption: $T_n(f)$ with $f \geq 0$ (spd).

Algorithmic proposals and Two Grid analysis

- Fiorentino, Serra-Capizzano, Calcolo '91 ($d = 1$), SISC '96 ($d = 2$) for τ (sine) and Toeplitz
- R. Chan, Chang, Sun [SISC '98], Cheng, Wang, Zhang, Tablino-Possio, Huckle, Staudacher, Fischer, Bolten, ...

V-cycle optimality

- Arico', D., Serra-Capizzano, SIMAX '04 ($d = 1$)
- Arico', D., NM '07 ($d > 1$)

Multigrid components

Smoother

Classical iterative method (damped Jacobi, ...)

Galerkin approach

For every level i

- $A_{i+1} = P_i A_i P_i^T$ (coarse matrix)

Projector

- To apply recursively the MGM, we need to maintain the same structure at every level (Toeplitz, circulant, ...).
- For every structure associated to the known BCs, we have preserving-structure projectors

$$P_i = K_{N_i} \mathcal{A}_{N_i}(p_i), \qquad \mathcal{A} \in \{T, C, \dots\}$$

- $K_{N_i} \in \mathbb{R}^{N_{i+1} \times N_i}$, $N_{i+1} < N_i$, is the preserving structure cutting matrix.
- p_i is the symbol associated to the projector and responsible of the subspace where the system is projected.

Optimality of the Multigrid in a Toeplitz setting

$T_n(f^{(q)})$ coefficient matrix, $p^{(q)}$ projector symbol satisfying the optimality conditions (d = dimensionality)

$$f^{(q)}(x) = \sum_{i=1}^{d} \left[2 - 2\cos(x_i)\right]^q \quad \text{e} \quad p^{(q)}(x) = \prod_{j=1}^{d} \left[2 + 2\cos(x_j)\right]^q$$

$d = 1$				$d = 2$			
n	# iterations			$n_1 \cdot n_2$	# iterations		
	$f^{(1)}$	$f^{(2)}$	$f^{(3)}$		$f^{(1)}$	$f^{(2)}$	$f^{(3)}$
$2^7 - \xi$	9	41	53	$(2^6 - \xi)^2$	6	24	33
$2^8 - \xi$	9	44	54	$(2^7 - \xi)^2$	6	26	33
$2^9 - \xi$	10	47	54	$(2^8 - \xi)^2$	6	27	33
$2^{10} - \xi$	9	48	55	$(2^9 - \xi)^2$	6	29	33

$\xi = 2\lceil (q+1)/2 \rceil - 1,\ q = 1, 2, 3.$

Ruge - Stuben conditions for V-cycle convergence

$$\|S_i \mathbf{x}\|_{A_i}^2 \leqslant \|\mathbf{x}\|_{A_i}^2 - \alpha_i \|\mathbf{x}\|_{A_i^2}^2 \quad (\alpha_i > 0) \qquad \forall \mathbf{x} \in \mathbb{C}^{n_i}, \qquad (1)$$

$$\|CGC_i \mathbf{x}\|_{A_i}^2 \leqslant \beta_i \|\mathbf{x}\|_{A_i^2}^2 \qquad \forall \mathbf{x} \in \mathbb{C}^{n_i}. \qquad (2)$$

(1): smoothing property. (2): approximation property.

$$\Downarrow$$

$$\|MGM_0\|_A \leqslant \sqrt{1-\delta} < 1, \qquad \delta = \min_{0 \leqslant i \leqslant m-1}\{\alpha_i/\beta_i\}.$$

MGM_0 is the V-cycle iteration matrix obtained by fixing $MGM_m = O_{n_m \times n_m}$ and computing for $i = m-1, \ldots, 0$

$$MGM_i = S_i[I_{n_i} - P_i^T(I_{n_{i+1}} - MGM_{i+1})A_{i+1}^{-1}P_iA_i].$$

Ruge - Stuben conditions $\rightarrow$ symbol conditions

Smoothing property:

$S_i = I_{n_i} - \omega_i A_i$ (Richardson) $\rightarrow \alpha_i$ s.t. $\alpha_i \leqslant \omega_i(2 - \omega_i \|f_i\|_\infty)$
(best $\omega_i = 1/\|f_i\|_\infty \Rightarrow \alpha_i \leq 1/\|f_i\|_\infty$).

Approximation property:

- $f(x) = f_0(x) = [1 - \cos(x)]^q \psi_0(x)$, with ψ_0 being a positive trigonometric polynomial (original symbol).
- $p(x) = \sqrt{2}\,[1 + \cos(x)]^q$, $i = 0, \ldots, m-1$ (chosen projector symbol).

$$\Downarrow$$

$f_i(x) = [1 - \cos(x)]^q \psi_i(x)$ (symbol at i-th recursion level), s.t.
$\{\psi_i\}$ is. a positive trigonometric polynomial.
(f_i vanishes only at zero with the same order $2q$ as f_0.)

Properties of $\{\psi_i\}$

$$\psi_{i+1}(x) = \left[\Phi_q(\psi_i)\right](x) = \frac{1}{2^{q+\frac{1}{2}}}\left[(p\,\psi_i)\left(\frac{x}{2}\right) + (p\,\psi_i)\left(\pi + \frac{x}{2}\right)\right].$$

- $\psi_i \in \mathbb{P}_q$ for i large enough (independently of ψ_0)
- $\psi_i(0) = \psi_0(0)$ (by induction);
- if $\psi_{i+1}(x) = 0$, $x > 0$, then $\psi_i(x/2) = 0$: $\psi_0(x)$ positive implies $\psi_i(x)$ positive for every i;
- if $\{\psi_i\}$ has a limit ψ^* and $\psi_0(x)$ positive, then $\psi^*(x)$ positive.

Defining $\mu_\infty(g) = \frac{\sup |g|}{\inf |g|}$, we have

$$\lim_{i\to\infty} \mu_\infty(\psi_i) = \mu_\infty(\psi^*) < +\infty.$$

Convergence of $\{\psi_i\} \Rightarrow$ optimality of V-cycle

Optimality of the V-cycle

$$\delta = \inf_n \min_{0 \leqslant i \leqslant \log_2(n)} \frac{\alpha_i}{\beta_i} > 0, \qquad \frac{\alpha_i}{\beta_i} \geqslant \frac{1}{2^q \mu_\infty^2(\psi_i)}.$$

If the sequence $\{\psi_i\}$ converges then

$$\|MGM_0\|_{A_0} \leq \sqrt{1 - \frac{1}{2^q \mu_\infty(\psi^*)^2}} < 1.$$

On the limit behavior of $\{\psi_i\}$ and Perron-Frobenius

$$\begin{cases} \psi \in \mathbb{P}_q \\ \psi(0) > 0 \end{cases} \quad \Rightarrow \quad \exists \psi^* \in \mathbb{P}_q : [\Phi_q]^i(\psi) \xrightarrow{\text{uniformly}} \psi^*.$$

As Φ_q is linear, by using the basis $\{e^{-\mathbf{i}qx}; \ldots; e^{\mathbf{i}qx}\}$ and by denoting by B_q the matrix representing Φ_q in such a basis, the preceding implication is equivalent to

$$\begin{cases} \mathbf{a} \in \mathbb{C}^{2q+1} \\ \sum_{j=1}^{2q+1} a_j > 0 \end{cases} \quad \Rightarrow \quad \exists \mathbf{a}^* \in \mathbb{C}^{2q+1} : \left[B_q\right]^i \mathbf{a} \longrightarrow \mathbf{a}^*.$$

B_q has one eigenvalue equal to 1 with algebraic multiplicity 1 and positive eigenvector $\mathbf{a}^*$, while all the other eigenvalues λ_i satisfy $|\lambda_i| < 1$ (via the Perron–Frobenius theorem).

V-cycle optimality: extensions

Some integral problems

If $f_0(\pi) = 0$ then $f_1(0) = 0$ (the projected problem is spectrally equivalent to a discretized differential problem) and the V-cycle is again optimal.

Multidimensional case ($d > 1$)

Differently to the algebra preconditioning, the previous optimality analysis can be extended to $d > 1$ [Arico', D., NM'07].

V-cycle optimality via symbols

Definition

The set of all corners of x is

$$\Omega(x) = \{\, y \,|\, y_j \in \{x_j,\, \pi + x_j\},\, j = 1, \ldots, d\}$$

and the set of the "mirror" points of x is $\mathcal{M}(x) = \Omega(x) \setminus \{x\}$.

Theorem (Aricò, D., NM '07)

Let x_0 be the unique zero of f_i in $[0,\pi]^d$, $\forall x \in [0,\pi]^d$, for $i = 0, \ldots, m-1$, we choose p_i such that

$$\limsup_{x \to x_0} \left| \frac{p_i(y)}{f_i(x)} \right| < +\infty, \qquad y \in \mathcal{M}(x),$$

where

$$0 < \sum_{y \in \Omega(x)} p_i^2(y).$$

Two Grid analysis via symbol vs Local Fourier Analysis

- In the case of constant coefficient PDEs and the Galerkin approach the two-grid LFA and the analysis by the symbol of Toeplitz matrices give the same results and comparable tools [D. NLAA, '10].
- The analysis by the symbol of Toeplitz matrices is more general since it includes also non differential problems, like for instance some integral problems.
- Future work: consider also the case of variable coefficients where available.

Deblurring problem

The restored signal/image $\mathbf{f}$ is obtained solving: (in some way by regularization ...)

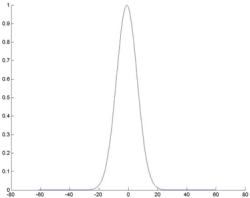
$$A\mathbf{f} = \mathbf{g} + \boldsymbol{\xi} = \hat{\mathbf{g}}$$

- $\mathbf{g}$ = blurred object,
- $\boldsymbol{\xi}$ = noise (random vector),
- A = (two-level) matrix with a Toeplitz-like structure depending on the point spread function (PSF) and the BCs.

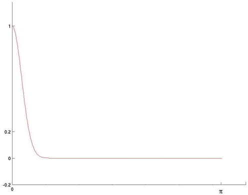
The PSF is the observation of a single point (e.g., a star in astronomy) that we assume shift invariant.

Generating function of the PSF

- The eigenvalues of $A(z)$ are about a uniform sampling of z.



PSF



Symbol $z(x)$

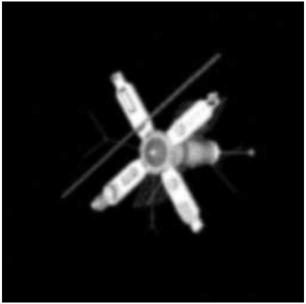
- The ill-conditioned subspace largely intersects high frequencies, while the well conditioned space essentially coincides with the low frequencies.

Without regularization

In order to obtain a fast convergence the multigrid projects in the high frequencies where the noise "lives" $\Longrightarrow$ noise explosion already at the first iteration.

Example

Symmetric blur + 2% of Gaussian noise



Blurred image with 2% of noise.

Restored image

Tikhonov regularization

- Tikhonov regularization:

$$\min_{\mathbf{f}\in\mathbb{R}^N} \left\{\| T_n(z)\mathbf{f} - \hat{\mathbf{g}}\|_2^2 + \mu\|\mathbf{f}\|_2^2\right\},$$

  which leads to

$$(T_n(z)^T T_n(z) + \mu I)\mathbf{f} = T_n(z)^T\hat{\mathbf{g}}.$$

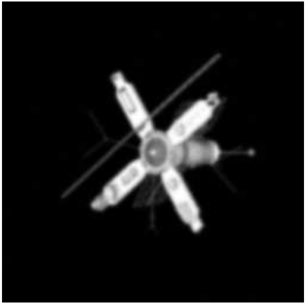
- Riley regularization ($z \geq 0$):

$$T_n(z+\theta)\mathbf{f} = \hat{\mathbf{g}}.$$

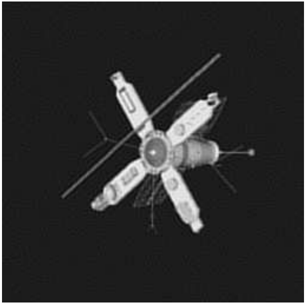
- Solved by multigrid for Toeplitz matrices ([D. NLAA '05]).

Riley linear system solved by multigrid

Observed image



Restored image



θ	10^{-2}	10^{-3}	10^{-4}
Linear interpolation	45	286	> 1000
Toeplitz (*)	49	73	73

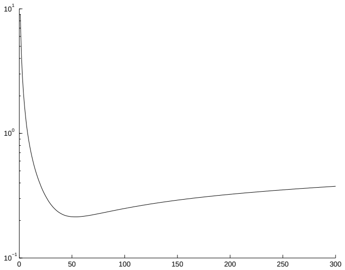
$$(*)\ p(x,y) = (2 - 2\cos(x))^2(2 - 2\cos(y))^2$$

Semi-convergence of iterative regularization methods

Let $\mathbf{f_o}$ the true object, it holds $A\mathbf{f_o} = \mathbf{g}$ but we "solve" $A\mathbf{f} = \hat{\mathbf{g}}$.

Some classical iterative methods (Landweber, CGLS, . . .):

- firstly reduce the algebraic error into the low frequencies (well-conditioned subspace).
- when they arrive to reduce the algebraic error into the high frequencies then the restoration error increases because of the noise.



Two-Level (TL) regularization

Idea: project into the low frequencies and then apply an iterative regularization method as smoother.

TL Algorithm

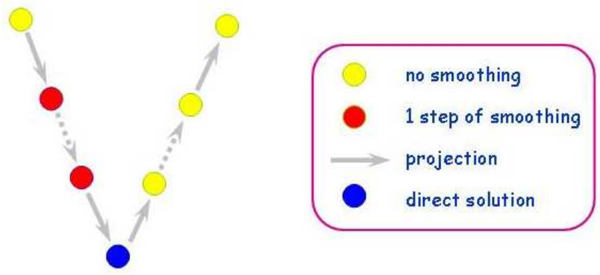
1. No smoothing at the finer level
2. At the coarser level to apply one step of the smoother instead of solving directly the linear system

Filter factor analysis

The filter factor analysis of the TL gives a condition to choose the projector s.t. we obtain an iterative regularization method.
[D., Serra-Capizzano, ETNA '07]

Multigrid regularization (applying recursively the TL)

V-cycle [D., Serra Capizzano, SISC '06]



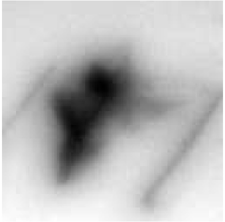
Using a larger number of recursive calls (e.g. W-cycle), the algorithm "works" more in the well-conditioned subspace, but it is more difficult to define an early stopping criterium.

Regularizing Multigrid

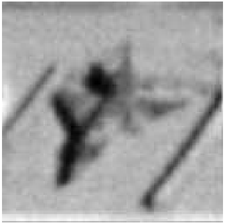
- Periodic BCs
- Gaussian PSF (A spd)
- Noise = 1%

True image

Internal part 128×128



Observed image



Restoration by MGM

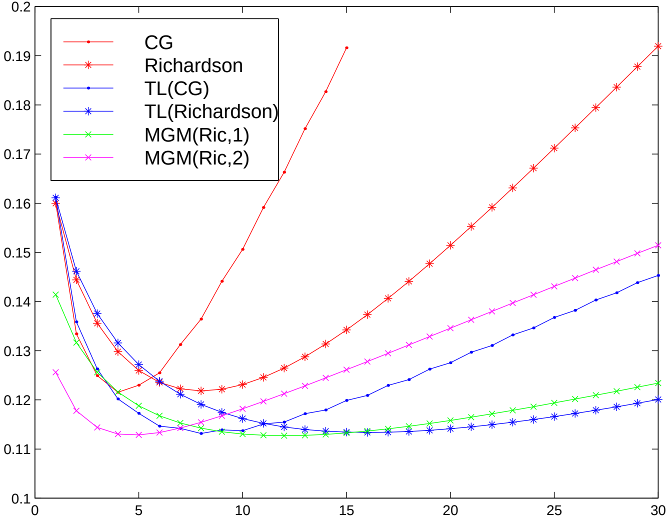
Restoration Error

$e_j = \|\mathbf{f_o} - \mathbf{f}^{(j)}\|_2 / \|\mathbf{f_o}\|_2$ error of reconstruction at the j-th iteration.

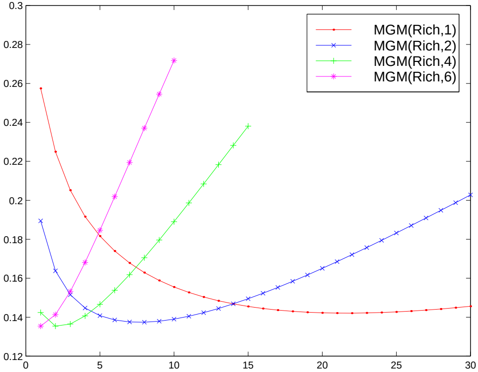
Minimal reconstruction error

Method	$\min_{j=1,\dots} (e_j)$	$\arg\min_{j=1,\dots} (e_j)$
CG	0.1215	4
Richardson	0.1218	8
TL(CG)	0.1132	8
TL(Rich)	0.1134	16
MGM(Rich, 1)	0.1127	12
MGM(Rich, 2)	0.1129	5
CGLS	0.1135	178
Landweber	0.1135	352

Relative error vs. number of iterations



The γ regularization



Restoration error
vs.
number of iterations

The computational cost increases with γ but not so much (e.g. $\gamma = 8 \Rightarrow O(N^{1.5})$ where $N = n^2$).

Conclusions

- Multigrid overcomes the curse of dimensionality.
- Exploiting the algebraic vs spectral structure of a problem can accelerate substantially the computation or improve the regularization property of classical iterative methods without spoiling the precision.

Future planes:

- Extend the range of application of such a mathematical technology to complex problems as those typically considered in the Engineering setting (GLT?).
- This afternoon: Session C3 “Toeplitz and LFA”.