Noise revealing in Golub-Kahan bidiagonalization as a mean of regularization in discrete inverse problems

Iveta Hnětynková, Marie Kubínová

Charles University in Prague & Institute of Computer Science AS CR

SNA'14 - Nymburk
January 2014

Outline

Discrete inverse problems

Noise revealing in Golub-Kahan iterative bidiagonalization

Denoising based on noise revealing

Conclusion

Outline

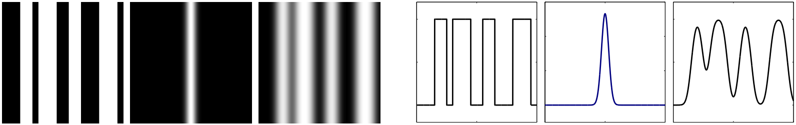
Fredholm integral equations of the first kind

Given the *continuous smooth kernel* $K(s,t)$ and the (measured) *data* $g(s)$, the aim is to find the (source) function $f(t)$ such that

$$g(s) = \int_I K(s,t) f(t) \mathrm{d}t.$$

Fredholm integral has smoothing property, i.e. high frequency components in g are dampened compared to f.

Example: Barcode reading



sharp barcode | Gaussian blur | blurred barcode | sharp barcode | Gaussian blur | blurred barcode

Discretization

Consider a discretization of the integral equation in the form of linear inverse problem

$$Ax \approx b, \quad b = b^{\text{exact}} + b^{\text{noise}}, \quad A \in \mathbb{R}^{m \times n}, \quad x \in \mathbb{R}^n,$$

where vector b^{noise}

- is an unknown perturbation representing rounding and discretization error, and/or noise with physical sources,
- resembles white noise, i.e., it has flat frequency characteristics,
- $\|b^{\text{noise}}\| \ll \|b^{\text{exact}}\|$.

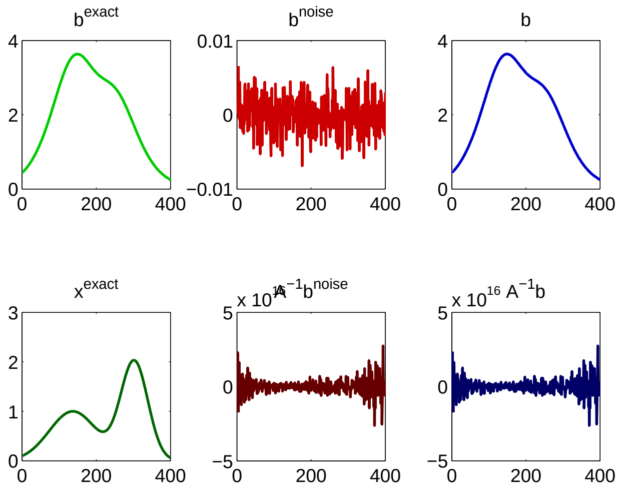
Aim is to approximate $x^{\text{exact}} \equiv A^\dagger b^{\text{exact}}$.

Remark

- A dampens high frequencies
- $A^\dagger$ amplifies high frequencies

Noise amplification

`shaw(400)`[1]



[1] Hansen: Regularization Tools

Noise amplification

The components of the "naive" solution

$$x^{\text{naive}} \equiv A^{\dagger} b = \underbrace{\sum_{j=1}^{l} \frac{u_j^T b^{\text{exact}}}{\sigma_j} v_j}_{x^{\text{exact}}} + \underbrace{\sum_{j=1}^{l} \frac{u_j^T b^{\text{noise}}}{\sigma_j} v_j}_{\text{amplified noise}}$$

$$+ \underbrace{\sum_{j=l+1}^{N} \frac{u_j^T b^{\text{exact}}}{\sigma_j} v_j}_{x^{\text{exact}}} + \underbrace{\sum_{j=l+1}^{N} \frac{u_j^T b^{\text{noise}}}{\sigma_j} v_j}_{\text{amplified noise}}$$

corresponding to small σ_j's are dominated by amplified noise.

Regularization is used to suppress the effect of amplified noise while extracting as much information about the solution as possible.

Overview of regularization methods

Spectral filtering (e.g., truncated SVD, Tikhonov): suitable for solving small ill-posed problems.

Projection methods (e.g., LSQR): suitable for solving large ill-posed problems. The size of projection space represents a regularization parameter [Björck - 88].

Hybrid methods: combination of outer iterative regularization with a spectral filtering of the projected small problem, see e.g., [Chung, Nagy, O'Leary - 08], [Kilmer, Hansen, Español - 06], [Kilmer, O'Leary - 01], [O'Leary, Simmons - 81].

Denoising

Suppose we have an estimate of b^{noise} available. Subtract the estimate $\tilde{b}^{\text{noise}}$ and solve transformed system

$$Ax \approx b^{\text{denoised}}, \quad \text{where} \quad b^{\text{denoised}} = b - \tilde{b}^{\text{noise}}.$$

In practice, the aim is to

- decrease the relative noise level in the right-hand side

and/or

- make the spectral properties of noise more favorable – dampen high frequencies.

Outline

Discrete inverse problems

Noise revealing in Golub-Kahan iterative bidiagonalization

Denoising based on noise revealing

Conclusion

Basic algorithm

Golub-Kahan iterative bidiagonalization (GK)

input: A, b;
define: $w_0 \equiv 0$, $s_1 \equiv b/\beta_1$, where $\beta_1 \equiv \|b\|$;
for $k = 1, 2, \ldots$

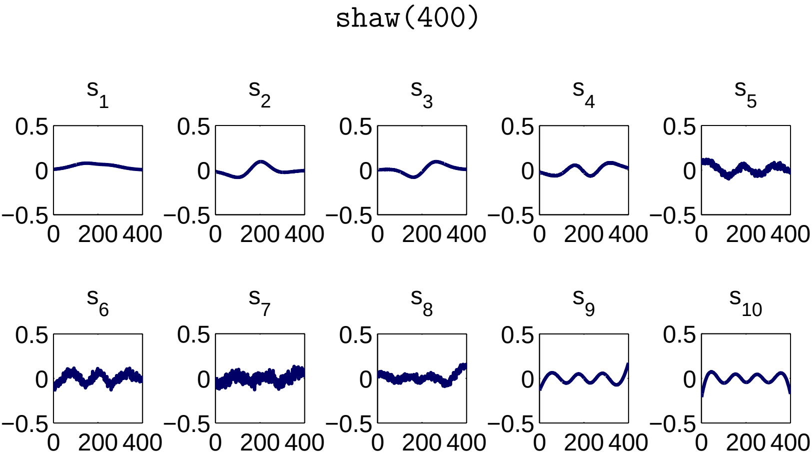
$$\alpha_k w_k = A^T s_k - \beta_k w_{k-1}\,, \qquad \|w_k\| = 1\,,$$
$$\beta_{k+1} s_{k+1} = A w_k - \alpha_k s_k\,, \qquad \|s_{k+1}\| = 1\,,$$

until $\alpha_k = 0$ or $\beta_{k+1} = 0$, or until $k = n$.

After k steps, GK produces:

- orthonormal vectors $s_1, \ldots, s_{k+1}$,
- orthonormal vectors $w_1, \ldots, w_k$,
- normalization coefficients $\alpha_1, \ldots, \alpha_k$ and $\beta_1, \ldots, \beta_{k+1}$.

Noise revealing in Golub-Kahan bidiagonalization



[Hnětynková, Plešinger, Strakoš - 09]

Noise propagation in Golub-Kahan bidiagonalization

Define $s_1^{\text{exact}} \equiv b^{\text{exact}}/\beta_1$, $s_1^{\text{noise}} \equiv b^{\text{noise}}/\beta_1$, and for $k = 1, 2, \ldots$

$$\begin{aligned} \alpha_k w_k &= \quad A^T s_k - \beta_k w_{k-1}\,, \\ \beta_{k+1} s_{k+1}^{\text{exact}} &\equiv \quad A w_k - \alpha_k s_k^{\text{exact}} \\ \beta_{k+1} s_{k+1}^{\text{noise}} &\equiv \qquad\quad - \alpha_k s_k^{\text{noise}} \end{aligned}$$

The troublesome high-frequency noise is confined to s_k^{noise}. Thus

$$s_{k+1}^{\text{noise}} = -\frac{\alpha_k}{\beta_{k+1}} s_k^{\text{noise}} = (-1)^k \prod_{j=1}^{k} \frac{\alpha_j}{\beta_{j+1}} s_1^{\text{noise}} = (-1)^k \prod_{j=1}^{k} \frac{\alpha_j}{\beta_{j+1}} \frac{b^{\text{noise}}}{\|b\|},$$

i.e., the study of white noise propagation reduces to the study of the cumulative amplification ratio

$$\rho_k^{-1} \equiv \prod_{j=1}^{k} \frac{\alpha_j}{\beta_{j+1}}\,.$$

[Hnětynková, Plešinger, Strakoš - 09]

Properties of the amplification ratio

GK for spectral components

$$\alpha_k(V^T w_k) = \Sigma\,(U^T s_k) - \beta_k(V^T w_{k-1})\,,$$
$$\beta_{k+1}(U^T s_{k+1}) = \Sigma\,(V^T w_k) - \alpha_k(U^T s_k)\,.$$

Vectors $U^T s_k$ and $V^T w_k$ exhibit dominance in the same components, therefore the orthogonality between s_{k+1} and s_k cannot be achieved without significant cancellation, and $\beta_{k+1} \ll \alpha_k$. Since the dominance in $U^T s_k$ and $V^T w_{k-1}$ is shifted by one component, $\alpha_k \approx \beta_k$.

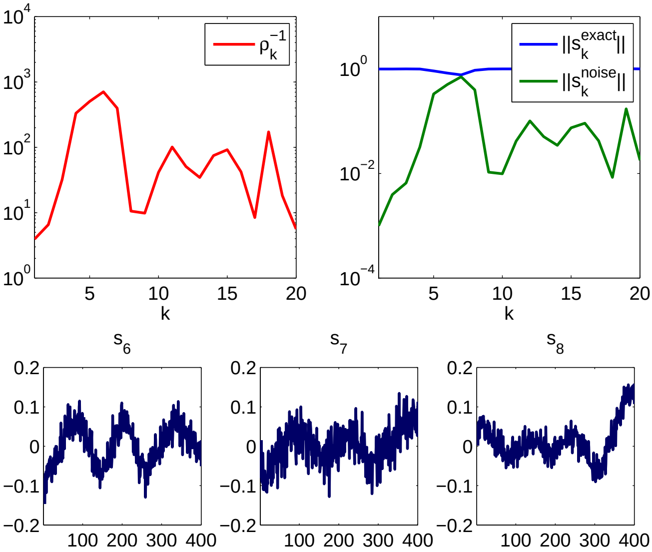
Therefore, the cumulative amplification ratio $\rho_k^{-1} \equiv \prod_{j=1}^{k} \frac{\alpha_j}{\beta_{j+1}}$ on average rapidly grows until it reaches the noise revealing iteration

$$k_{\text{noise}} \equiv \underset{k}{\operatorname{argmax}}\, \rho_k^{-1}.$$

At this step s_{k+1} is dominated by s_{k+1}^{noise}.

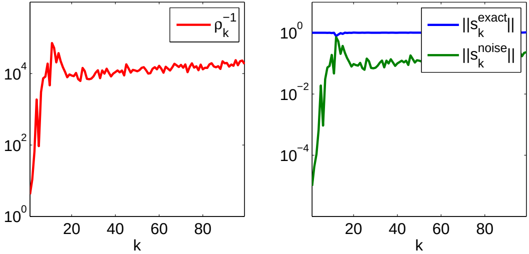
Noise revealing

shaw(400)

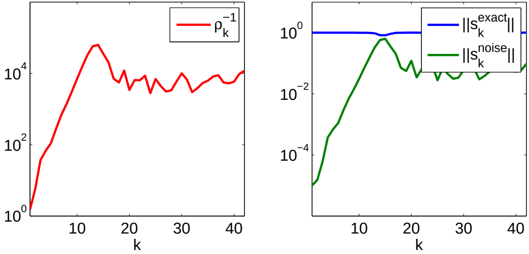


Examples

phillips(400)



i_laplace(400,1)



Outline

Denoising by noise revealing

Recall

$$s_{k+1}^{\text{noise}} = (-1)^k \rho_k^{-1} \frac{b^{\text{noise}}}{\|b\|},$$

which yields

$$b^{\text{noise}} = \|b\|(-1)^k \rho_k s_{k+1}^{\text{noise}}.$$

At step $k = k_{\text{noise}}$, s_{k+1} is dominated by s_{k+1}^{noise}. Take

$$b^{\text{noise}} \approx \tilde{b}^{\text{noise}} \equiv \|b\|(-1)^k \rho_k s_{k+1},$$

i.e., estimate noise by properly scaled noise revealing left bidiagonalization vector.

Define the denoised right-hand side

$$b^{\text{denoised}} \equiv b - \tilde{b}^{\text{noise}}.$$

Properties of denoised right-hand side

It was shown in [Kubínová - 13] that from

$$b^{\text{denoised}} = b - \tilde{b}^{\text{noise}}, \quad \tilde{b}^{\text{noise}} = \|b\|(-1)^k \rho_k s_{k+1}$$

and from

$$b^{\text{exact}} = \|b\| \left[s_1 - s_1^{\text{noise}}\right] = \|b\| \left[s_1 - (-1)^k \rho_k s_{k+1}^{\text{noise}}\right],$$

we get

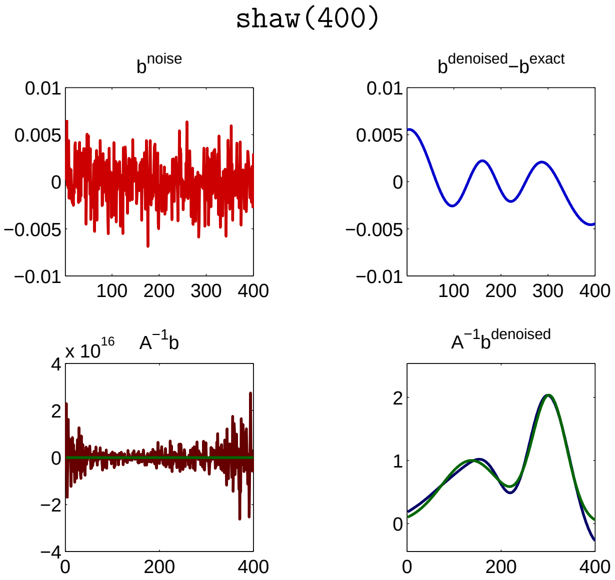
$$b^{\text{denoised}} - b^{\text{exact}} = -\|b\|(-1)^k \rho_k s_{k+1}^{\text{exact}}.$$

Consequently

- remaining noise is smooth - high frequency part of noise is not present;
- new relative noise level is

$$\frac{\|b^{\text{denoised}} - b^{\text{exact}}\|}{\|b^{\text{exact}}\|} \approx \rho_k \|s_{k+1}^{\text{exact}}\| \approx \sqrt{\rho_k^2 - \delta_{\text{noise}}^2}.$$

Regularization effect of denoising



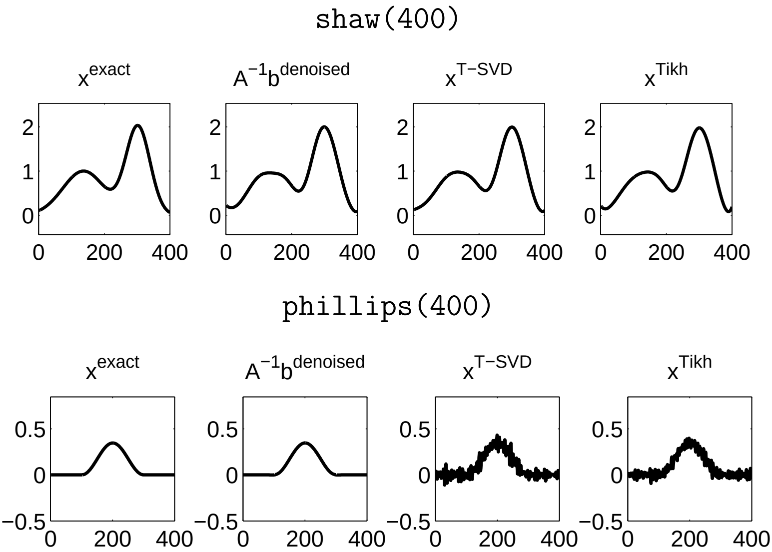
Quantitative properties

shaw(400)				
δ_{noise}	1e-02	1e-04	1e-06	1e-08
$k_{noise}+1$	5	8	10	13
$\rho_{k_{noise}}$	1.09e-02	1.05e-04	1.32e-06	1.15e-08
$\frac{\|b^{\text{denoised}}-b^{\text{exact}}\|}{\|b^{\text{exact}}\|}$	4.57e-03	3.67e-05	8.73e-07	6.30e-09

i_laplace(400,1)				
δ_{noise}	1e-02	1e-04	1e-06	1e-08
$k_{noise}+1$	6	12	17	23
$\rho_{k_{noise}}$	1.89e-02	1.62e-04	1.61e-06	1.57e-08
$\frac{\|b^{\text{denoised}}-b^{\text{exact}}\|}{\|b^{\text{exact}}\|}$	1.62e-02	1.29e-04	1.28e-06	1.26e-08

phillips(400)				
δ_{noise}	1e-02	1e-04	1e-06	1e-08
$k_{noise}+1$	5	9	16	32
$\rho_{k_{noise}}$	1.39e-02	1.44e-04	1.41e-06	2.15e-08
$\frac{\|b^{\text{denoised}}-b^{\text{exact}}\|}{\|b^{\text{exact}}\|}$	9.96e-03	1.06e-04	1.09e-06	2.00e-08

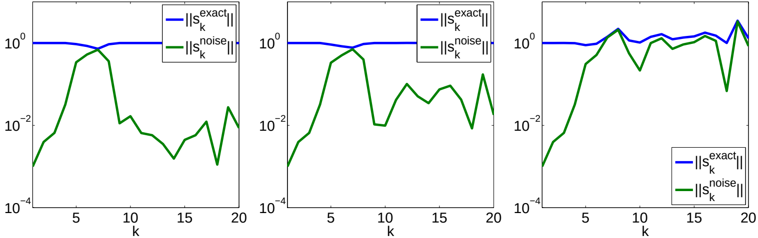
Comparison with other methods



Denoising for colored noise

Generally:

High-frequency noise is considered the more convenient alternative – distinguish easier between low-frequency exact data and noise.



high-frequency   white   low-frequency

Observation:

For low-frequency noise, $\|s_k^{\text{exact}}\| + \|s_k^{\text{noise}}\| \not\approx 1$.

Outline

Conclusion

Method at a glance

- Golub-Kahan iterative bidiagonalization
- find noise revealing iteration $\operatorname{argmax}_k \rho_k^{-1} = \operatorname{argmax}_k \prod_{j=1}^{k} \frac{\alpha_j}{\beta_{j+1}}$
- subtract properly scaled noise revealing left bidiagonalization vector from b, i.e., $b^{\text{denoised}} = b - \|b\|(-1)^k \rho_k s_{k+1}$

Remarks

- the method is extremely cheap
- loss of orthogonality delays noise revealing

Open questions

- how to solve $Ax \approx b^{\text{denoised}}$, further regularization
- relation to other methods

Thank you for your attention.