# Noise propagation on the Golub-Kahan iterative bidiagonalization process

Juan Durazo, Brendan Horan, Arthur Mitrano

School of Mathematical and Statistical Sciences
Arizona State University

April, 2013

# Outline

# $Ax = b$

- We want to solve the problem:

$$Ax = b, \quad A \in \mathbb{R}^{n \times n}, \quad b = b^{exact} + b^{noise}.$$

- $b^{noise}$ is a white noise vector with unknown noise level $\delta_{noise}$.
- Information on how the noise propagates can be used to develop a stopping criteria for hybrid methods (iterate and regularize).
- The Golub-Kahan iterative bidiagonalization process provides a way to analyze how the noise spreads into the core problem.

# Golub-Kahan Iterative Bidiagonalization process

- Given $w_0 = 0$, $s_1 = b/\beta_1$, where $\beta_1 = \|b\| \neq 0$:

$$\begin{aligned} \alpha_j w_j &= A^T s_j - \beta_j w_{j-1}, & \|w_j\| &= 1, \\ \beta_{j+1} s_{j+1} &= A w_j - \alpha_j s_j, & \|s_{j+1}\| &= 1 \end{aligned}$$

  until $\alpha_j = 0$ or $\beta_{j+1} = 0$, or we reach the dimensionality of the problem.
- The vectors $s_k$ and $w_k$ are orthonormal.

# Matrix form

- Let $S_k = [s_1, \ldots, s_k]$ and $W_k = [w_1, \ldots, w_k]$, then:
$$A^T S_k = W_k L_k^T, \quad AW_k = [S_k, s_{k+1}]L_{k+},$$
where
$$L_k = \begin{bmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \beta_k & \alpha_k \end{bmatrix}, \quad L_{k+} = \begin{bmatrix} L_k \\ \beta_{k+1} e_k^T \end{bmatrix}$$
- This bidiagonal decomposition can be used to calculate the SVD of $A$ in a stable way. $A = S_k L_k W_k^T = S_k U \Sigma V^T W_k^T$ $= U_k \Sigma V_k^T$, with $U_k$ and $V_k$ unitary.

# Lanczos tridiagonalization of $AA^T$

- Using the starting vector $s_1 = b/\beta_1$, $\beta_1 = \|b\|$, yields in $k$ steps:

$$AA^T S_k = S_k T_k + \alpha_k \beta_{k+1} s_{k+1} e_k^T,$$

and

$$T_k = L_k L_k^T = \begin{bmatrix} \alpha_1^2 & \alpha_1\beta_2 & & \\ \alpha_1\beta_2 & \alpha_2^2+\beta_2^2 & \ddots & \\ & \ddots & \ddots & \alpha_{k-1}\beta_k \\ & & \alpha_{k-1}\beta_k & \alpha_k^2+\beta_k^2 \end{bmatrix}$$

- The matrix $L_k$ from the GKIB is a Cholesky factor of the matrix $T_k$.

# The core problem

- Consider $x_k = W_k y_k$ as an approximation for the solution of $Ax_k \approx b$.
- Looking to the residual:

$$\begin{aligned} r_k = b - AW_k y_k &= S_{k+1}(\beta_1 e_1 - L_{k+} y_k) \\ &= S_k(\beta_1 e_1 - L_k y_k) - (\beta_{k+1} e_k^T y_k) s_{k+1}, \end{aligned}$$

we get two subproblems:

1. $L_k y_k = \beta_1 e_1$, for $r_k$ orthogonal to $S_k$ (CGME);
2. $y_k = \min_y \|L_{k+} y - \beta_1 e_1\|$, if we minimize $r_k$ (LSQR or CGLS).

# Summarizing

- The bidiagonalization concentrates the useful information of the main problem on its bidiagonal block.
- In presence of noise, those subproblems might be polluted by the noise.
- An investigation on how the noise propagates might aid on solving the ill-posed problem. (When to stop the iteration procedure).
- We will use two main approaches: one looks to the behavior of the high frequencies of the vectors $s_k$ and the other its normalized cumulative periodogram.

- The test problem utilized for the numerical experiments was `shaw(400)`, which satisfy the *discrete Picard condition* on average.
- We used the following noise levels $\delta_{noise} = 10^{-14}, 10^{-8}, 10^{-4}$. Where

$$\delta_{noise} = \frac{\|b^{noise}\|}{\|b^{exact}\|}.$$

# Picard plot

- $u_j$ are the left singular vectors of $A$.

# Spectral coefficients

# Left singular vectors

# Analizing the spectral coefficients

- Using the SVD of $A$ and the Lanczos tridiagonalization of $AA^T$:

$$\Sigma^2(U^T S_k) = (U^T S_k)(L_k L_k^T) + \alpha_k \beta_{k+1}(U^T s_{k+1})e_k^T.$$

- Setting $k = 1$ we can see how the noise spreads to $s_2$ (looking to the last column):

$$\alpha_1 \beta_2 (U^T s_2) = (\Sigma^2 - \alpha_1^2 I) U^T s_1$$

- The matrix acts like a filter on the lower frequencies

$$\frac{\sigma_i}{\alpha_1 \beta_2} - \frac{\alpha_1}{\beta_2}$$

- $\alpha_1/\beta_2$ is likely to be larger than 1.
- The noise level on the high frequencies will get larger on $s_2$

# Generalizing

- For the general case we can write:

$$U^T s_{k+1} = \phi_k(\Sigma^2) U^T s_1$$

- $\phi_k(\lambda)$ is the Lanczos polynomial with root $(\theta_l^{(k)})^2$, $l = 1, \ldots, k$ (Ritz values).
- The large Ritz values $(\theta_k^{(k)})^2$, $(\theta_k^{(k)})^2, \ldots$ closely approximate the singular values $\sigma_1^2, \sigma_2^2, \ldots$.
- Note that the constant term

$$\phi_k(0) = \prod_{j=1}^{k} \frac{\alpha_j}{\beta_{j+1}} = \rho_k^{-1}$$

- $\phi_k$ acts like a filter on low frequencies and the constant term $\phi_k(0) = \rho_k^{-1}$ causes the amplification of the high frequency noise present in the noisy vector $s_1$.

# Signal and noise spaces

- $span\{u_1, \ldots, u_{k+1}\}$ is the signal subspace.
- $span\{u_{k+2}, \ldots, u_n\}$ is the noise subspace.

# Behavior of $s_k^{exact}$ and $s_k^{noise}$

- To further illustrate the noise amplification we look to $s_k^{exact}$ and $s_k^{noise}$. Let $s_1 = s_1^{exact} + s_1^{noise}$ and

$$\beta_{k+1} s_{k+1}^{exact} = Aw_k - \alpha_k s_k^{exact},$$
$$\beta_{k+1} s_{k+1}^{noise} = -\alpha_k s_k^{noise},$$
$$s_{k+1} = s_{k+1}^{exact} + s_{k+1}^{exact}, \quad \beta_{k+1} s_{k+1} = Aw_k - \alpha_k s_k.$$

- They aren't the true exact and noise data, but they give good approximations to the euclidean norm. See [1].
- Note that,

$$s_{k+1}^{noise} = -\frac{\alpha_k}{\beta_{k+1}} s_k^{noise} = (-1)^k \rho_k^{-1} s_1^{noise}.$$

# Normalized Cumulative Periodogram (NCP)

- Definition:

$$\mathbf{c}_j = |dft(\mathbf{y})_j|^2, \quad z_j = \frac{\sum_{i=1}^{j} \mathbf{c}_i}{\sum_{i=1}^{q} \mathbf{c}_i}, \quad j = 1, \ldots, q.$$

- Kolmogorov-Smirnoff at 5% significance level: the NCP curve must lie between the limits $\pm 1.35 q^{-1/2}$ of the straight line (diagonal).
- There are other kind of measures for testing if a distribution is white-noise like, e.g., *total deviation*.

# NCP plot

# White-noise measures

# Comparing the methods to find $k_{noise}$

| problem | `shaw(400)` | | | |
|---|---|---|---|---|
| $\delta_{noise}$ | $10^{-14}$ | $10^{-8}$ | $10^{-4}$ | $10^{-2}$ |
| $k_{noise}$ (GKIB) | 17 | 11 | 7 | 3 |
| $k_{noise}$ (NCP) | 27 | 26 | 8 | 23 |
| $\tilde{\delta}_{noise}$ (GKIB) | $5.9E-15$ | $1.2E-8$ | $5.3E-5$ | $1.6E-2$ |
| $\tilde{\delta}_{noise}$ (NCP) | $1.5E-15$ | $8.1E-9$ | $3.7E-4$ | $3.6E-3$ |

- $\tilde{\delta}_{noise}$ indicates the estimated noise level.

# Resolution matrix of the GKIB process

- $R^\sharp = W_k(L_k^T L_k)^{-1} L_k^T S_k^T A$.
- Show the resolution matrix in Matlab.

Hnetynkova, I. and Plesinger, M.. *The regularizing effect of the Golub-Kahan iterative bidiagonalization and revealing the noise level in the data.* BIT Numer Math (2009) 49: 669-696.