Vector estimates for the action of matrix functions on vectors

Paraskevi Roupa

Joint work with: Marilena Mitrouli

Department of Mathematics

University of Athens, Greece

Numerical Linear Algebra with Applications (NL2A) **October 27, 2016**

Outline of the talk

1. Introduction
2. Vector estimates for f(A)b
3. Estimates for matrix functionals
4. Numerical examples
5. Concluding remarks

Introduction

Given:

- A: a **diagonalizable** matrix of order p
- b: a vector of order p and
- f : an analytic function defined on the spectrum of the matrix A

Task:

- Approximation of the **action** of f(A) **on a vector** b, i.e. **f(A)b**, without computing f(A).

Applications

It is not necessary either to estimate the whole matrix f(A) or it is not feasible to compute f(A).

* lattice quantum chromodynamics computations in chemistry and physics
* numerical solution of stochastic differential equations
* sampling from a Gaussian process distribution

Ref:

- J. Chen, M. Anitescu, Y. Saad, Computing f(A)b via least squares polynomial approximations, SIAM, 33 (2011), 195-222.
- P. I. Davies, N. J. Higham, Computing f(A)b for Matrix Functions f, Vol. 47 (2005) of the series Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin, 15-24.

Vector estimates for f(A)b

* **Eigendecomposition** of A:

$$A = Q\,\Lambda\,Q^{-1}$$

where

- **Λ:** pxp diagonal matrix which contains the eigenvalues λ_i of A and
- **Q:** pxp matrix which contains the corresponding linearly independent eigenvectors of A.

* We can write:

$$f(A) = Q\,f(\Lambda)\,Q^{-1} = \sum_{i=1}^{p} f(\lambda_i) q_i \hat{q}_i^T$$

* We define the **vector-moments**:

$$d_r = A^r b \quad \text{and} \quad d_f = f(A)b.$$

* We can write:

$$d_r = A^r b = \sum_{i=1}^{p} \lambda_i{}^r q_i(\hat{q}_i{}^T, b) = \sum_{i=1}^{p} \lambda_i{}^r a_i q_i$$

and

$$d_f = f(A)\, b = \sum_{i=1}^{p} f(\lambda_i) q_i(\hat{q}_i{}^T, b) = \sum_{i=1}^{p} f(\lambda_i) a_i q_i$$

where:

$$\alpha_i = (\hat{q}_i{}^T, b).$$

One – term vector estimates

$$d_0 = A^0\, b \approx \alpha\, q$$
$$d_1 = A^1\, b \approx l\, \alpha\, q = l\, d_0$$
$$d_2 = A^2\, b \approx l^2\, \alpha\, q = l^2\, d_0$$

Family of **one-term** vector estimates for **f(A)b**:

$$\varphi_N(i) = f\,(d_0(i)^{N(i)-1}\, d_1(i)^{1-2N(i)}\, d_2(i)^{N(i)})\, d_0(i),\ \ N(i) \in \mathbb{C},$$

$$i = 1, 2\,, \ldots\,, p.$$

Proposition

The family of one-term vector estimates satisfies the relation

$$\varphi_N(i) = f\,(\rho(i)^{N(i)}\,\frac{d_1(i)}{d_0(i)}\,)\;d_0(i), \quad i = 1, 2\,, \ldots\,, p,$$

where $\rho(i) = \dfrac{d_0(i)d_2(i)}{d_1(i)^2}$.

Lemma

Let $A \in \mathbb{R}^{p\text{x}p}$ be a diagonalizable matrix and f an invertible function. There exists a vector $N_{opt} \in \mathbb{C}^p$ which the i-th element is given by

$$N_{opt}(i) = \frac{\log\left(f^{-1}\left(\frac{d_f(i)}{d_0(i)}\right)\frac{d_0(i)}{d_1(i)}\right)}{\log(\rho(i))}, \ \rho(i) = \frac{d_0(i)d_2(i)}{d_1(i)^2} \neq 1, \ i = 1, 2, \ldots, p,$$

such that $\varphi_{N_{opt}}$ gives the exact value of f(A)b.

Lemma

Let A ∈ $\mathbb{R}^{pxp}$ be a diagonalizable matrix and f an increasing function. If $d_0(i) > 0$, $d_1(i) > 0$ and $\rho(i) > 1$ then a bound for the optimal value $N_{opt}(i)$ is

$$N_{opt}(i) \leq \frac{\log\left(f^{-1}\left(k(Q) f(\rho(\mathrm{A}))\frac{||b||}{d_0(i)}\right)\frac{d_0(i)}{d_1(i)}\right)}{\log(\rho(i))}, \quad i=1, 2, \ldots, p,$$

where k(Q) is the condition number of the matrix of the eigenvectors Q and ρ(A) is the spectral radius of A.

Two – term vector estimates

$$d_0 \approx \alpha_1 q_1 + \alpha_2 q_2$$
$$d_1 \approx l_1 \alpha_1 q_1 + l_2 \alpha_2 q_2$$
$$d_2 \approx l_1^2 \alpha_1 q_1 + l_2^2 \alpha_2 q_2$$

The family of **two-term** vector estimates for **f(A)b** satisfies the relation:

$$\hat{\varphi}_{n,\kappa}(i) = f\,(l_1(i))\,\alpha_1 q_1(i) + f\,(l_2(i))\,\alpha_2 q_2(i), \quad i = 1, 2, \ldots, p,$$

where:

$$\alpha_1\, q_1(i) = \frac{1}{l_2(i)-l_1(i)}\,(l_2(i)\, d_0(i) - d_1(i)), \quad l_1(i) \neq l_2(i),$$

$$\alpha_2\, q_2(i) = \frac{1}{l_2(i)-l_1(i)}\,(d_1(i) - l_1(i)\, d_0(i)\ ), \quad l_1(i) \neq l_2(i),$$

$$l_{1,2}(i) = \frac{r(i) \pm \sqrt{r(i)^2 - 4q(i)}}{2}$$

$$r(i) = \frac{d_{n-1}(i)\, d_{n+2+k}(i) - d_{n+1}(i)\, d_{n+k}(i)}{d_{n-1}(i)\, d_{n+1+k}(i) - d_n(i)\, d_{n+k}(i)}, \quad n, k \in \mathbb{Z}$$

$$q(i) = \frac{d_n(i)\, d_{n+2+k}(i) - d_{n+1}(i)\, d_{n+1+k}(i)}{d_{n-1}(i)\, d_{n+1+k}(i) - d_n(i)\, d_{n+k}(i)}, \quad n, k \in \mathbb{Z}.$$

Estimates for matrix functionals

- Let A be a **diagonalizable** matrix of order p and x, y be vectors of order p.
- Estimates for matrix functionals of the form:

$$\mathbf{x^* f(A) y}$$

- We define the moments:

$$c_r(x,y) = (x, A^r y) \quad \text{and} \quad c_f(x,y) = (x, f(A)y).$$

Ref: P. Fika, M. Mitrouli, Estimation of the bilinear form $y^* f(A)x$ for Hermitian matrices, Linear Algebra Appl., 502, pp. 140-158, 2015.

* **Family** of **one-term** estimates:

$$e_\nu = f\,(c_0^{\nu-1}\, c_1^{1-2\nu}\, c_2^{\nu})\, c_0, \;\; \nu \in \mathbb{C}.$$

* The **optimal value** of ν is given by the type

$$\nu_{\mathrm{opt}} = \frac{\log\left(f^{-1}\left(\frac{c_f}{c_0}\right)\frac{c_0}{c_1}\right)}{\log(\rho)},\ \rho = \frac{c_0 c_2}{c_1{}^2} \neq 1.$$

* **Family** of **two-term** estimates

$$\hat{e}_{n,\kappa} = f\,(l_1)\,\alpha_1 q_1 + f\,(l_2)\,\alpha_2 q_2$$

where:

$$\alpha_1\, q_1 = \frac{1}{l_2 - l_1}\,(l_2\, c_0 - c_1), \quad l_1 \neq l_2,$$

$$\alpha_2\, q_2 = \frac{1}{l_2 - l_1}\,(c_1 - l_1\, c_0), \quad l_1 \neq l_2,$$

$$l_{1,2} = \frac{r \pm \sqrt{r^2 - 4q}}{2}$$

$$r = \frac{c_{n-1}\, c_{n+2+k} - c_{n+1}\, c_{n+k}}{c_{n-1}\, c_{n+1+k} - c_n\, c_{n+k}}, \quad q = \frac{c_n\, c_{n+2+k} - c_{n+1}\, c_{n+1+k}}{c_{n-1}\, c_{n+1+k} - c_n\, c_{n+k}}, \quad n, k \in \mathbb{Z}$$

Connection between f(A)b and $x^T f(A)y$

- We have $x = e_i$ and $y = b$.

- Moments:

  $c_r(x,y) = (d_r)_i$ and $c_f(x,y) = (d_f)_i$, $i = 1, 2, \ldots, p$.

- We can write $f(A)b = \begin{bmatrix} < e_1, f(A)b > \\ < e_2, f(A)b > \\ \vdots \\ < e_p, f(A)b > \end{bmatrix}$.

Numerical examples

We estimate the quantity **f(A)b**, for various functions f, matrices A and vectors b.

Complexity: $O(sp^2)$, $s \leq 8$

Ref:

- The University of Florida Sparse Matrix Collection, *http://www.cise.ufl.edu/research/sparse/matrices/*.
- The Matlab gallery, http://www.mathworks.com/help/matlab/ref/gallery.html.

Example 1: ***Diagonalizable matrices***

We test the matrix A = **dw256B** of order **p = 512**.

- diagonalizable with positive eigenvalues
- well conditioned (k(A) = 3.7328)

We estimate the quantity $\mathbf{A^{1/2}b}$ with

- b is drawn from the uniform distribution
- b is drawn from the normal distribution.

(n,k)	relative error [b = rand(p,1)]	relative error [b = randn(p,1)]
(1,0)	9.9679e-4	7.7755e-4
(1,1)	1.9075e-3	1.1618e-3
(1,3)	3.6696e-3	1.3892e-3
(1,-2)	3.6671e-4	8.3626e-4
(0,4)	3.5097e-3	4.4683e-3

Estimating $\boldsymbol{A^{1/2}b}$ by using the family of ***two-term*** *estimates for various values of the parameters* $\boldsymbol{n}$ *and* $\boldsymbol{k}$*.*

* Estimating **f(A)b** for various A, b, f.

  Two-term vector estimates with n=1 and k=0.

matrix A	**vector b**	**function f(A)**	**relative error**
ex1 (p=216)	randn(216,1)	$A^{1/2}$	3.6759e-2
ex1 (p=216)	rand(216,1)	exp(A)	9.8744e-7
parter (p=800)	$b_i = \cos(i)$	exp(A)	4.0454e-2
parter (p=800)	e_5	$A^{1/2}$	4.6376e-2
rand(300)/100	randn(300,1)	exp(A)	5.2986e-3

*Estimating **f(A)b** by using the family of **two-term estimates** with **n=1** and **k=0**.*

Example 2: ***Symmetric positive definite matrices***

We consider the **Covariance** matrix:

- symmetric positive definite
- entries: $\alpha_{ij} = \begin{cases} 1 + i, & i = j \\ \frac{1}{|i-j|^5}, & i \neq j \end{cases}$
- order p=700

Estimating **log(A)b** with:

- b = randn(700,1) and
- $b_i = \cot(i)$, i = 1, 2, …, 700.

vector b	(n,k)	relative error
randn(700,1)	(1,0)	1.8048e-4
	(1,3)	8.1004e-4
	(1,5)	8.6469e-4
b_i = cot(i)	(1,0)	6.1381e-7
	(3,5)	6.7791e-5
	(-4,1)	3.1369e-5

*Estimating **log(A)b** for various values of the parameters n and k.*

Concluding remarks

- **Families of vector estimates** for f(A)b were produced with complexity of order $O(p^2)$.

- The **presented numerical results** show the **satisfactory behavior** of the two-term vector estimates.

  In applications where a high accuracy is not required, the derived estimates are efficient and easily applicable.

References

* C. Brezinski, P. Fika, M. Mitrouli, Moments of a linear operator on a Hilbert space, with applications to the trace of the inverse of matrices and the solution of equations, Numer. Linear Algebra Appl., 19, pp. 937-953, 2012.
* C. Brezinski, P. Fika, M. Mitrouli, Estimations of the trace of powers of positive self-adjoint operators by extrapolation of the moments, Elec. Trans. Numer. Anal., 39, pp. 144-155, 2012.
* P. Fika, M. Mitrouli, P. Roupa, Estimates for the bilinear form $x^T A^{-1} y$ with applications to linear algebra problems, Elec. Trans. Numer. Anal., 43, pp. 70-89, 2014.
* G.H. Golub, G. Meurant, Matrices, Moments and Quadrature with Applications, Princeton University Press, 2010.
* N. J. Higham, Matrix Functions, Functions of Matrices: Theory and Computation, SIAM, 2008.
* N. J. Higham, A. H. Al-Mohy, Computing Matrix Functions, Acta Numerica, pp. 159-208, 2010.

Thank you!