Electronic Transactions on Numerical Analysis.
Volume 39, pp. 144-155, 2012.
Copyright © 2012, Kent State University.
ISSN 1068-9613.

ETNA
Kent State University
http://etna.math.kent.edu

ESTIMATIONS OF THE TRACE OF POWERS OF POSITIVE SELF-ADJOINT OPERATORS BY EXTRAPOLATION OF THE MOMENTS*

CLAUDE BREZINSKI†, PARASKEVI FIKA‡, AND MARILENA MITROULI‡

Abstract. Let A be a positive self-adjoint linear operator on a real separable Hilbert space H. Our aim is to build estimates of the trace of A^q, for $q \in \mathbb{R}$. These estimates are obtained by extrapolation of the moments of A. Applications of the matrix case are discussed, and numerical results are given.

Key words. Trace, positive self-adjoint linear operator, symmetric matrix, matrix powers, matrix moments, extrapolation.

AMS subject classifications. 65F15, 65F30, 65B05, 65C05, 65J10, 15A18, 15A45.

1. Introduction. Let A be a positive self-adjoint linear operator from H to H, where H is a real separable Hilbert space with inner product denoted by $(\cdot,\cdot)$. Our aim is to build estimates of the trace of A^q, for $q \in \mathbb{R}$. These estimates are obtained by extrapolation of the integer moments $(z, A^n z)$ of A, for $n \in \mathbb{N}$. A similar procedure was first introduced in [3] for estimating the Euclidean norm of the error when solving a system of linear equations, which corresponds to $q = -2$. The case $q = -1$, which leads to estimates of the trace of the inverse of a matrix, was studied in [4]; on this problem, see [10].

Let us mention that, when only positive powers of A are used, the Hilbert space H could be infinite dimensional, while, for negative powers of A, it is always assumed to be a finite dimensional one, and, obviously, A is also assumed to be invertible. With this convention, the two cases could be treated simultaneously. Moreover, since some of our results are valid in the infinite dimensional case, the mathematical concepts needed are given in their full generality in Section 2.

Traces of powers of matrices arise in several fields of mathematics. More specifically

- *Network analysis: triangle counting in a graph.* When analyzing a complex network, an important problem is to compute the total number of triangles of a connected simple graph. This number is equal to $\mathrm{Tr}(A^3)/6$ where A is the adjacency matrix of the graph [1]. For many networks, even if the matrix A is sparse, A^3 can be rather dense and, thus, it is not possible to compute this trace directly.
- *Number theory and combinatorics: Euler congruence.* Traces of powers of integer matrices are connected with the Euler congruence [18], an important phenomenon in mathematics, stating that

$$\mathrm{Tr}(A^{p^r}) \equiv \mathrm{Tr}(A^{p^{r-1}}) \pmod{p^r},$$

  for all integer matrices A, all primes p, and all $r \in \mathbb{N}$. The diversity of proofs of the Euler congruence indicates its universality and its role in different branches of mathematics.
- *Statistics: specification of classical optimality criteria.* In optimal design of experiments [15], the ultimate purpose of any optimality criterion is to measure the

*Received January 12, 2012. Accepted April 11, 2012. Published online May 7, 2012. Recommended by L. Reichel.

†Laboratoire Paul Painlevé, UMR CNRS 8524, UFR de Mathématiques Pures et Appliquées, Université des Sciences et Technologies de Lille, 59655–Villeneuve d'Ascq cedex, France (Claude.Brezinski@univ-lille1.fr).

‡University of Athens, Department of Mathematics, Panepistimiopolis, 15784 Athens, Greece ({pfika,mmitroul}@math.uoa.gr).

"largeness" of a nonnegative definite matrix C of dimension s. One of the most prominent criteria is the average variance criterion $\Phi_{-1}(C) = (\mathrm{Tr}(C^{-1})/s)^{-1}$, if C is positive definite. Invariance under reparametrization loses its appeal if the parameters of interest have a definite physical meaning. The above average variance criterion provides a reasonable alternative. More generally, for positive definite matrices C, the matrix mean Φ_p can be defined for every real parameter p by $\Phi_p(C) = (\mathrm{Tr}(C^p)/s)^{1/p}$, for $p \neq 0, \pm\infty$.

- *Dynamical systems: determination of their invariants.* The invariants of dynamical systems are described in terms of the traces of powers of integer matrices, for example in studying the Lefschetz numbers [18].
- *p-adic analysis: determination of the Witt vector.* For any integer matrix M and any prime number p, the entries of the unique Witt vector consisting of p-adic integers are expressed from the traces of powers of the integer matrix M [18].
- *Matrix theory: extremal eigenvalues.* There are many applications in matrix theory and numerical linear algebra. For example, in order to obtain approximations of the smallest and the largest eigenvalues of a symmetric matrix A, a procedure based on estimates of the trace of A^n and A^{-n}, $n \in \mathbb{N}$, was proposed in [14].
- *Differential equations: solution of Lyapunov matrix equation.* These equations can be solved by using matrix polynomials and characteristic polynomials where the computation of the traces of matrix powers are needed [8].

The computation of the trace of matrix powers has received much attention. In [8], an algorithm for computing $\mathrm{Tr}(A^k)$, $k \in \mathbb{Z}$, is proposed, when A is a lower Hessenberg matrix with a unit codiagonal. In [7], a symbolic calculation of the trace of powers of tridiagonal matrices is presented.

The mathematical tools needed are described in Section 2. The extrapolation procedure for estimating the moments of an operator is presented in Section 3. Estimates for the trace of powers of a matrix are derived in Section 4. Numerical results are given in Section 5. Concluding remarks end the paper.

2. The mathematical background. Let A be a compact positive self-adjoint operator over a separable infinite dimensional Hilbert space H. The eigenvalues λ_k of A are real and positive, and it exists an orthonormal basis of H consisting of its corresponding eigenelements $\{u_k\}$ (the operator A can be diagonalized by an orthonormal set of eigenvectors).

We have

$$Au_k = \lambda_k u_k, \quad k = 1, 2, \ldots,$$

and also

$$\lambda_k = (u_k, Au_k) = (Au_k, Au_k)^{1/2} \quad \text{and} \quad (u_k, Au_n) = 0, \quad n \neq k.$$

A bounded linear operator A over a separable Hilbert space H is said to be in the *trace class* if the sum

$$\sum_k ((A^*A)^{1/2} u_k, u_k)$$

is finite, where $\{u_k\}$ is any orthonormal basis [16, p. 32]. In this case, the trace of A is defined by the absolutely convergent sum

$$\mathrm{Tr}(A) = \sum_k (Au_k, u_k),$$

ETNA
Kent State University
http://etna.math.kent.edu

which is independent of the choice of the basis. Lidskii's theorem [13] states that, if A is a positive compact operator and if $\{u_k\}$ is any orthonormal basis of H, its trace is equal to the sum of its non-zero eigenvalues, each of them enumerated with its algebraic multiplicity.

We remind the canonical form of a compact self-adjoint operator on a Hilbert space

$$\forall z \in H, \quad Az = \sum_k \lambda_k (z, u_k) u_k.$$

3. The extrapolation procedure for the moments. For $q \in \mathbb{R}$, the canonical form of the powers of the operator A will be defined by

$$A^q z = \sum_k \lambda_k^q (z, u_k) u_k,$$

and its moments by

$$c_q(z) = (z, A^q z) = \sum_k \lambda_k^q \alpha_k^2(z), \tag{3.1}$$

where $\alpha_k(z) = (z, u_k)$.

We will now provide estimates of the trace of A^q. These estimates are based on the integer moments of A which are defined by (3.1), with $q = n \in \mathbb{N}$. Obviously, in practice, only the moments with $n \geq 0$ can be computed.

The moments $c_q(z)$ defined in (3.1) are given by a sum. Starting from some moments $c_n(z)$ with a nonnegative integer index n, we want to estimate the moments $c_q(z)$ for any fixed index $q \in \mathbb{R}$. For this purpose, we will interpolate these $c_n(z)$'s by a conveniently chosen function defined by keeping only one or two terms in the summation (3.1), and then extrapolate this function at the point q. This idea was introduced in [3] for estimating the norm of the error in the solution of a system of linear algebraic equations, and it was used in [4] for estimating the trace of the inverse of a matrix.

3.1. One-term estimates. We want to estimate $c_q(z)$, $q \in \mathbb{R}$, by keeping only one term in formula (3.1), that is by a function of the form

$$c_q(z) \simeq e_q(z) = s^q a^2(z).$$

Thus, knowing the values of $c_0(z)$ and $c_1(z)$, the interpolation conditions

$$c_0(z) \simeq e_0(z) = a^2(z), \qquad c_1(z) \simeq e_1(z) = s a^2(z).$$

give us the 2 unknowns s and $a(z)$, and we have the

PROPOSITION 3.1. *The moment $c_q(z)$ can be estimated by the direct one-term formula*

$$c_q(z) \simeq e_q(z) = \frac{c_1^q(z)}{c_0^{q-1}(z)}, \qquad q \in \mathbb{R}. \tag{3.2}$$

REMARK 3.2. The estimate $e_q(z)$ of formula (3.2) is real since A is a positive self-adjoint operator and, thus, $c_1(z)$ is positive.

REMARK 3.3. For $q = -1$, we have $c_{-1}(z) \simeq e_{-1}(z) = c_0^2(z)/c_1(z)$, which is the one-term estimate of $c_{-1}(z)$ given in [4], which leads to the estimate of the trace of the inverse of a matrix.

ETNA
Kent State University
http://etna.math.kent.edu

Let us now assume that A^{-1} exists, and let κ be the Euclidean condition number of A defined by $\kappa = \|A\| \cdot \|A^{-1}\|$. We have the

PROPOSITION 3.4. *If A is self-adjoint positive definite, then, for any vector z, the one-term estimate $e_n(z)$ given by (3.2) satisfies the following inequalities for $n \in \mathbb{Z}$, $n \neq 0$,*

$$e_n(z) \leq c_n(z) \leq \left(\frac{(1+\kappa)^2}{4\kappa} \right)^{2^d - 1} e_n(z),$$

where

$$d = \begin{cases} n-1, & n > 1, \\ |n|, & n < 0, \; n = 1. \end{cases}$$

Proof. We have, by an inequality given in [5, Theorem 4],

$$c_i^n(z) \leq c_{in}(z)\, c_0^{n-1}(z), \quad i = 0, \pm 1, \pm 2, \ldots, \quad n \geq 0.$$

Thus, for $i = 1$, it follows

$$c_1^n(z)/c_0^{n-1}(z) \leq c_n(z). \tag{3.3}$$

Moreover, from [5, Theorem 1], it holds

$$c_{n+1}(z) \leq \frac{(1+\kappa)^2}{4\kappa} \cdot \frac{c_n^2(z)}{c_{n-1}(z)}. \tag{3.4}$$

We will prove by induction that

$$c_n(z) \leq \left(\frac{(1+\kappa)^2}{4\kappa} \right)^{2^{n-1}-1} \frac{c_1^n(z)}{c_0^{n-1}(z)}, \quad n > 1. \tag{3.5}$$

This inequality is true when $n = 2$, since it is the inequality (3.4) for $n = 1$. Assume that (3.5) holds for $n \in \mathbb{N}$, and let us prove that it still holds for $n+1$. From (3.4) we have,

$$c_{n+1}(z) \leq \frac{(1+\kappa)^2}{4\kappa} \cdot \frac{c_n^2(z)}{c_{n-1}(z)},$$

where $c_n^2(z)$ can be upper bounded from (3.5) since all the quantities $c_i(z)$ are positive and the inequality can be squared. As $c_{n-1}(z)$ can be lower bounded from (3.3) by replacing n by $n-1$, we have

$$c_{n+1}(z) \leq \frac{(1+\kappa)^2}{4\kappa} \frac{((1+\kappa)^2/4\kappa)^{2(2^{n-1}-1)} c_1^{2n}(z)/c_0^{2n-2}(z)}{c_1^{n-1}(z)/c_0^{n-2}(z)},$$

and the result immediately follows, because

$$\frac{(1+\kappa)^2}{4\kappa} \cdot \frac{((1+\kappa)^2/4\kappa)^{2(2^{n-1}-1)} c_1^{2n}(z)/c_0^{2n-2}(z)}{c_1^{n-1}(z)/c_0^{n-2}(z)} = \left(\frac{(1+\kappa)^2}{4\kappa} \right)^{2^n - 1} \frac{c_1^{n+1}(z)}{c_0^n(z)}.$$

The inequality for $n < 0$ and $n = 1$ can be proved in a similar way. □

ETNA
Kent State University
http://etna.math.kent.edu

3.2. Two-term estimates. We want now to estimate $c_q(z)$, $q \in \mathbb{R}$, by keeping two terms in formula (3.1), that is by a function of the form

$$c_q(z) \simeq e_q(z) = s_1^q a_1^2(z) + s_2^q a_2^2(z). \tag{3.6}$$

The four unknowns $s_1, s_2, a_1^2(z)$, and $a_2^2(z)$, will be computed by imposing the interpolation conditions

$$c_n(z) = s_1^n a_1^2(z) + s_2^n a_2^2(z), \tag{3.7}$$

for four different values of the integer n, namely $n = 0, \ldots, 3$. Writing this relation for $n = 0$ and 1 gives a system of two equations in the two unknowns $a_1^2(z)$ and $a_2^2(z)$ if s_1 and s_2 are known. But the interpolation conditions (3.7) mean, in fact, the $c_n(z)$'s satisfy the difference equation of order 2

$$c_{n+2}(z) - sc_{n+1}(z) + pc_n(z) = 0,$$

where $s = s_1 + s_2$ and $p = s_1 s_2$. Using this relation for $n = 0$ and 1 gives s and p. Then, we obtain

$$\begin{aligned} s &= \frac{c_0(z)c_3(z) - c_1(z)c_2(z)}{c_0(z)c_2(z) - c_1^2(z)}, \quad & a_1^2(z) &= \frac{c_0(z)s_2 - c_1(z)}{s_2 - s_1}, \\ p &= \frac{c_1(z)c_3(z) - c_2^2(z)}{c_0(z)c_2(z) - c_1^2(z)}, & a_2^2(z) &= \frac{c_1(z) - c_0(z)s_1}{s_2 - s_1}, \end{aligned} \tag{3.8}$$

and the estimate (3.6) for $e_q(z)$ follows with $s_{1,2} = (s \pm \sqrt{s^2 - 4p})/2$. Thus, we have the following result

PROPOSITION 3.5. *The moment $c_q(z)$ can be estimated by the* direct two-term formula

$$c_q(z) \simeq e_q(z) = s_1^q a_1^2(z) + s_2^q a_2^2(z), \qquad q \in \mathbb{R},$$

where s_1, s_2, a_1^2, and a_2^2, are given by the formulae (3.8).

REMARK 3.6. The direct two-term estimate $e_q(z)$ of this Proposition is real if $q \in \mathbb{R}$. Indeed, if $s^2 - 4p < 0$, s_1 and s_2 are complex conjugate, and it is easy to see that a_1^2 and a_2^2 are also complex conjugate and, thus, $e_q(z) \in \mathbb{R}$. In case $s^2 - 4p \geq 0$, then s_1 and s_2 are real. Also s_1 and s_2 are positive, because $p = s_1 s_2$, as defined by relations (3.8), is positive; this can be deduced from considering the inequality given in [5, Theorem 1], $c_{i+1}^2(z) \leq c_i(z)c_{i+2}(z)$, $i = 0, \pm 1, \pm 2, \ldots$, for $i = 0$ and $i = 1$. Moreover, $s = s_1 + s_2$, as defined by relations (3.8), is positive; this follows from the inequality given in [5, Theorem 4], $c_1(z)/c_0(z) \leq c_3(z)/c_2(z)$. Thus, $e_q(z) \in \mathbb{R}$.

Looking at (3.6), the $e_n(z)$'s also satisfy the difference equation

$$e_{n+2}(z) - se_{n+1}(z) + pe_n(z) = 0.$$

Thus, using again the formulae (3.8) for s and p, the $e_n(z)$'s can be recursively computed, for integer values of the index, and we obtain the

PROPOSITION 3.7. *The moment $c_n(z)$ can be estimated by the* forward iterative formula

$$e_n(z) = se_{n-1}(z) - pe_{n-2}(z), \qquad n = 2, 3, \ldots,$$

or by the backward iterative formula

$$e_n(z) = (se_{n+1}(z) - e_{n+2}(z))/p, \qquad n = -1, -2, \ldots,$$

ETNA
Kent State University
http://etna.math.kent.edu

with $e_0(z) = (z, z)$ *and* $e_1(z) = (z, Az)$.

REMARK 3.8. Taking $n = -1$ in the backward iterative formula, and using s and p from (3.8), we obtain the formula

$$c_{-1}(z) \simeq e_{-1}(z) = \frac{{c_1}^3(z) + {c_0}^2(z)c_3(z) - 2c_0(z)c_1(z)c_2(z)}{c_1(z)c_3(z) - {c_2}^2(z)},$$

which leads to the two-term estimate given in [4] for the trace of the inverse of a matrix.

The results of this Section will now be used for estimating the trace of the powers of a symmetric positive definite matrix. Although not indicated, all sums run from 1 up to the dimension of the matrix.

4. Estimates for traces.

Let A be a symmetric positive definite matrix, and let $\{\lambda_k\}$ denote its eigenvalues. For $q \in \mathbb{R}$, A^q is also symmetric positive definite, and it holds [11]

$$\mathrm{Tr}(A^q) = \sum_k \lambda_k^q.$$

Our estimates for the trace of A^q are based on the following result proved by Hutchinson [12]; see also [2] and [10, p. 170].

PROPOSITION 4.1. *Let* $M = (m_{ij})$ *be a symmetric matrix of dimension* p *with* $\mathrm{Tr}(M) \neq 0$. *Let* X *be a discrete random variable taking the values* 1 *and* -1 *with the equal probability* 0.5, *and let* x *be a vector of* p *independent samples from* X*; for simplicity, we write, in this case,* $x \in X^p$. *Then* (x, Mx) *is an unbiased estimator of* $\mathrm{Tr}(M)$, *and it holds*

$$E((x, Mx)) = \mathrm{Tr}(M)$$

and

$$\mathrm{Var}((x, Mx)) = 2\sum_{i\neq j} m_{ij}^2,$$

where $E(\cdot)$ *and* $\mathrm{Var}(\cdot)$ *denote the expected value and the variance respectively.*

This Proposition tells us that $\mathrm{Tr}(A^q) = E((x, A^q x)) = E(c_q(x))$, for $x \in X^p$. Thus, for the one-term estimates (3.2), it immediately follows from Proposition 3.4.

COROLLARY 4.2. *If the matrix* A *is symmetric positive definite, then, for the direct one-term estimates* (3.2), *we have the bounds*

$$E(e_n(x)) \leq \mathrm{Tr}(A^n) \leq \left(\frac{(1+\kappa)^2}{4\kappa}\right)^{2^d - 1} E(e_n(x)), \quad n \in \mathbb{Z},\ n \neq 0,$$

where

$$d = \begin{cases} n-1, & n > 1, \\ |n|, & n < 0,\ n = 1. \end{cases}$$

Notice that, if A is orthogonal, then $\kappa(A) = 1$, and $\mathrm{Tr}(A^n) = E(e_n(x))$.

For the direct one-term formula (3.2) of Proposition 3.1, the expectation $E(e_n(x))$ appearing in the bounds of the inequality of Corollary 4.2 is given by the formula

$$E(e_n(x)) = E(c_1^n(x)/c_0^{n-1}(x)) = E(c_1^n(x))/N^{n-1}, \quad x \in X^p. \tag{4.1}$$

ETNA
Kent State University
http://etna.math.kent.edu

In practice, the computation of $c_1^n(x)$ needs quite tedious algebraic developments. Indeed, $c_1(x) = \sum_{i,j} a_{ij}\xi_i\xi_j$, where $x = (\xi_1, \ldots, \xi_p)^T$. Its nth power has first to be expanded, noticing that each term of $c_1^n(x)$ has the form $C\xi_1^{n_1} \cdots \xi_k^{n_k}$, with $\xi_i \neq \xi_j$ for $i \neq j$, $n_1 + \cdots + n_k = 2n$, and where C is a coefficient which is the product of n elements of the matrix A. Then, $E(c_1^n(x))$ has to be computed taking into account that, since the ξ_i's are independent random variables, the expectation of their product is equal to the product of the expectations, and that, for all i, $E(\xi_i^m)$ is equal to 1 if m is even and to 0 if m is odd. Unless a general formula could be obtained, the use of a computer algebra software is required for $n > 2$. A closed formula and numerical results for $n = 2$ are given in Example 5.4.

When $q \in \mathbb{R}$, estimates of $\mathrm{Tr}(A^q)$ can be obtained by realizing N experiments, and then computing the mean value of the quantities $e_q(x_i)$, for $x_i \in X^p$. We set

$$t_q = \frac{1}{N}\sum_{i=1}^{N} e_q(x_i), \tag{4.2}$$

where the x_i's are realizations of $x \in X^p$. Thus, formula (3.2) gives us the direct one term trace estimates t_q,

$$t_q = \frac{1}{N}\sum_{i=1}^{N} c_1^q(x_i)/c_0^{q-1}(x_i), \quad q \in \mathbb{R},$$

while, from formula (3.6), we have the following direct two term trace estimates

$$t_q = \frac{1}{N}\sum_{i=1}^{N} s_1^q a_1^2(x_i) + s_2^q a_2^2(x_i), \quad q \in \mathbb{R},$$

together with (3.8). Similarly, estimates for the variance are given by

$$v_q = \frac{\sum_{i=1}^{N}(e_q(x_i) - t_q)^2}{N-1}. \tag{4.3}$$

Proposition 3.7 leads us to the following result

PROPOSITION 4.3. *For $n \geq 2$, estimates of $\mathrm{Tr}(A^q)$, $q \in \mathbb{R}$, are given by the forward iterative formula*

$$t_n = st_{n-1} - pt_{n-2},$$

with the initial values $t_1 = \mathrm{Tr}(A)$ and $t_0 = p$ (since $A^0 = I$). For $n < 0$, estimates of $\mathrm{Tr}(A^q)$, $q \in \mathbb{R}$, are given by the backward iterative formula

$$t_n = (st_{n+1} - t_{n+2})/p,$$

with the same initial values.

The following result specifies the confidence interval for the trace estimate t_q of the trace of powers of symmetric positive definite matrices. Let us remind that the amount of evidence required to accept that an event is unlikely to arise by chance is known as the *significance level*. The lower the significance level, the stronger the evidence. The choice of the level of significance is arbitrary, but for many applications, a value of 5% is chosen, for no better reason than that it is conventional.

ETNA
Kent State University
http://etna.math.kent.edu

Let $Z_{a/2}$ be the upper $a/2$ percentage point of the normal distribution $\mathcal{N}(0,1)$. Then, the following result is a classical one about the probability of having a good estimate.

PROPOSITION 4.4.

$$\Pr\left(\left|\frac{t_q - \mathrm{Tr}(A^q)}{\sqrt{\mathrm{Var}((x, A^q x))/N}}\right| < Z_{a/2}\right) = 1 - a,$$

where N is the number of trials, a is the significance level, and $Z_{a/2}$ the critical value of the standard normal distribution defined above.

For a significance level $a = 0.01$, we have $Z_{a/2} = 2.58$, and Proposition 4.4 gives us a confidence interval for $\mathrm{Tr}(A^q)$ with probability $100(1-a)\% = 99\%$. Thus, we expect, for any sample's size, the trace estimate t_q to be in this interval with a probability of 99%.

If, in Proposition 4.4, $\mathrm{Var}((x, A^q x))$ is replaced by v_q given by formula (4.3), an approximation of the confidence interval is obtained.

5. Numerical results.

Let us now give some numerical results for illustrating the trace estimates t_q. Each realization of e_q requires only few matrix-vector products and some inner products. For a real dense symmetric matrix A of dimension p, the one-term estimate e_q needs $\mathcal{O}(p^2)$ flops, whereas the two-term one requires $\mathcal{O}(2p^2)$ flops. In the case of a banded matrix the complexity is reduced. Specifically, the one-term estimate e_q requires $\mathcal{O}(mp)$ flops, where m is the bandwidth. The two-term estimate has twice this complexity. Obviously, the computation of t_q by (4.2) and of v_q by (4.3) needs N times these flops.

TABLE 5.1
Estimations of $\mathrm{Tr}(P^{3/2})$.

Dim.	Exact	1-term est.	rel1	conf. interval
100	2.461e2	2.454e2	2.628e-3	[2.420e2, 2.489e2]
200	4.923e2	4.878e2	9.275e-3	[4.823e2, 4.932e2]
500	1.231e3	1.211e3	1.647e-2	[1.202e3, 1.220e3]
1000	2.462e3	2.421e3	1.700e-2	[2.409e3, 2.432e3]

TABLE 5.2
Estimations of $\mathrm{Tr}(P^{3/2})$.

Dim.	Exact	2-term est.	rel2	conf. interval
100	2.461e2	2.465e2	1.769e-3	[2.426e2, 2.504e2]
200	4.923e2	4.923e2	1.350e-4	[4.869e2, 4.976e2]
500	1.231e3	1.231e3	4.156e-5	[1.223e3, 1.239e3]
1000	2.462e3	2.464e3	6.400e-4	[2.453e3, 2.475e3]

TABLE 5.3
Estimations of $\mathrm{Tr}(P^{1/2})$.

Dim.	Exact	1-term est.	rel1	conf. interval
100	1.332e2	1.340e2	6.505e-3	[1.332e2, 1.349e2]
200	2.663e2	2.676e2	4.603e-3	[2.664e2, 2.687e2]
500	6.657e2	6.698e2	6.160e-3	[6.682e2, 6.714e2]
1000	1.331e3	1.341e3	7.535e-3	[1.339e3, 1.344e3]

ETNA
Kent State University
http://etna.math.kent.edu

For the vectors $x_i \in X^p$, we used the uniform generator of random numbers between 0 and 1 of MATLAB. If the random number obtained is less or equal to 0.5, the corresponding component of x_i is set to -1; if it is greater than 0.5 and smaller or equal to 1, the corresponding component of x_i is set to $+1$.

The relative errors for the one-term and two-term estimates are denoted by *rel1* and *rel2*, respectively. The condition number of the matrix is denoted by *cond*. The confidence intervals for $\mathrm{Tr}(A^q)$ of Proposition 4.4 were obtained using the numerical variance v_q computed by (4.3).

All computations were performed in MATLAB, the test matrices were given by the *gallery* function, and we took $N = 50$. The so-called *exact* values reported in this section are those given by the function *trace* of MATLAB. The matrix powers A^q, if q is an integer, are computed by repeated multiplication. If the integer is negative, A is inverted first. For other values of q, the calculation involves eigenvalues and eigenvectors, such that if $[V, D] = \mathrm{eig}(A)$, then $A^q = V * D^q * \mathrm{inv}(V)$, where "eig" and "inv" are the MATLAB internal functions for the computation of the eigenvalues and the inverse of the matrix, respectively.

EXAMPLE 5.1 (the Prolate matrix). We consider the Prolate matrix P. It is a symmetric Toeplitz matrix depending on a parameter w. If $0 < w < 0.5$, it is positive definite, its eigenvalues are distinct, lie in (0,1], and tend to cluster around 0 and 1. It is ill-conditioned if w is close to 0. In our examples we take $w = 0.9$ for which $\mathrm{cond}(P) = 2$. In Tables 5.1 and 5.2, we give the results for $\mathrm{Tr}(P^{3/2})$ and, in Tables 5.3 and 5.4, those for $\mathrm{Tr}(P^{1/2})$.

TABLE 5.4
Estimations of $\mathrm{Tr}(P^{1/2})$.

Dim.	Exact	2-term est.	rel2	conf. interval
100	1.332e2	1.332e2	3.4839e-5	[1.325e2, 1.339e2]
200	2.663e2	2.662e2	3.0015e-4	[2.651e2, 2.674e2]
500	6.657e2	6.656e2	2.0770e-4	[6.634e2, 6.678e2]
1000	1.331e3	1.332e3	1.2375e-4	[1.329e3, 1.334e3]

Let us now see the behavior of our estimates for a higher power. The results for $\mathrm{Tr}(P^{12})$ are presented in Tables 5.5 and 5.6. For the power $-1/2$, we obtain the results of Table 5.7.

TABLE 5.5
Estimations of $\mathrm{Tr}(P^{12})$.

Dim.	Exact	1-term est.	rel1	conf. interval
100	3.219e5	1.266e5	6.065e-1	[1.098e5, 1.435e5]
200	6.490e5	2.485e5	6.171e-1	[2.250e5, 2.719e5]
500	1.631e6	5.957e5	6.348e-1	[5.640e5, 6.274e5]
1000	3.269e6	1.189e6	6.363e-1	[1.139e6, 1.239e6]

TABLE 5.6
Estimations of $\mathrm{Tr}(P^{12})$.

Dim.	Exact	2-term est.	rel2	conf. interval
100	3.219e5	3.220e5	2.716e-4	[3.114e5, 3.325e5]
200	6.490e5	6.492e5	3.231e-4	[6.366e5, 6.618e5]
500	1.631e6	1.629e6	1.183e-3	[1.609e6, 1.650e6]
1000	3.269e6	3.263e6	1.803e-3	[3.233e6, 3.293e6]

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 5.7
Estimating $\mathrm{Tr}(P^{-1/2})$.

Dim.	Exact	2-term est.	rel2	conf. interval
100	7.647e1	7.619e1	3.710e-3	[7.568e1,7.670e1]
200	1.530e2	1.531e2	3.926e-4	[1.522e2, 1.539e2]
500	3.827e2	3.826e2	2.642e-4	[3.811e2, 3.841e2]
1000	7.656e2	7.645e2	1.314e-3	[7.626e2, 7.665e2]

EXAMPLE 5.2 (dense matrices). We consider the *Parter* matrix P whose elements are $p_{ij} = 1/(i-j+0.5)$. P is a Cauchy and a Toeplitz matrix. We set $A = P^T P$. The condition number for the matrix A of dimension $100,\ 200,\ 500,\ 1000$ has the values $10.997,\ 12.892,\ 15.638,\ 17.898$, respectively. The results for the trace of A^{15} are given in Tables 5.8 and 5.9.

TABLE 5.8
Estimations of $\mathrm{Tr}(A^{15})$.

Dim.	Exact	1-term est.	rel1	conf. interval
100	7.934e16	7.348e16	7.395e-2	[7.128e16, 7.568e16]
200	1.612e17	1.530e17	5.052e-2	[1.507e17, 1.554e17]
500	4.072e17	3.984e17	2.161e-2	[3.954e17, 4.014e17]
1000	8.176e17	8.067e17	1.339e-2	[8.031e17, 8.102e17]

TABLE 5.9
Estimations of $\mathrm{Tr}(A^{15})$.

Dim.	Exact	2-term est.	rel2	conf. interval
100	7.934e16	7.937e16	3.652e-4	[7.884e16, 7.991e16]
200	1.612e17	1.611e17	5.764e-4	[1.603e17, 1.619e17]
500	4.072e17	4.071e17	4.145e-4	[4.064e17, 4.078e17]
1000	8.176e17	8.177e17	1.463e-4	[8.170e17, 8.185e17]

EXAMPLE 5.3 (ill-conditioned sparse matrices). The following ill-conditioned matrices, denoted by B and whose dimensions are indicated into parenthesis in the Tables, come from the *Florida Sparse Matrix Collection* [9]. We tested two matrices for the power $3/2$ in Table 5.10, and five of them for the power 3 in Table 5.11. These matrices appear in stiffness problems, except the matrix *journals* in Table 5.11 which corresponds to an undirected weighted graph.

EXAMPLE 5.4 (bounds of Corollary 4.2). Let us give some numerical results for illustrating the bounds given in Corollary 4.2 for the trace of A^2. In that case, $E(e_2(x))$ can be exactly computed by Formula (4.1) with $n = 2$, where

$$E(c_1^2(x)) = 4\sum_{i<j} a_{ij}^2 + 2\sum_{i<j} a_{ii}a_{jj} + \sum_{i} a_{ii}^2.$$

This formula was already given by Hutchinson [12]. Its interest lies in the fact that it only implies the knowledge of A, and that A^2 does not have to be computed. Let us remind that, for a symmetric matrix, the trace of A^2 is the square of its Frobenius norm. The Frobenius norm

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 5.10
Estimations of $\mathrm{Tr}(B^{3/2})$.

Matrix	cond	Exact	2-term est.	rel2	conf. interval
bcsstk20(485)	7.48e12	8.610e24	8.664e24	6.199e-3	[7.943e24, 9.384e24]
bcsstk21(3600)	4.49e7	9.105e14	9.020e14	9.339e-3	[8.923e14, 9.117e14]

TABLE 5.11
Estimations of $\mathrm{Tr}(B^3)$.

Matrix	cond	Exact	2-term est.	rel2	conf. interval
bcsstk20(485)	7.483e12	9.145e48	9.226e48	8.763e-3	[8.261e48, 1.019e49]
bcsstk21(3600)	4.497e7	8.750e26	8.759e26	9.774e-4	[8.613e26, 8.905e26]
bcsstm06(420)	3.457e6	2.581e13	2.581e13	0	[2.581e13, 2.581e13]
bcsstm08(1074)	8.266e6	6.415e18	6.415e18	1.724e-14	[6.415e18, 6.415e18]
journals(124)	1.938e4	3.701e14	3.455e14	6.640e-2	[1.831e14, 5.079e14]

has application in obtaining lower and upper bounds for the Frobenius condition number on the cone of symmetric and positive definite matrices [6, 17].

As an illustration, we first consider the Kac-Murdock-Szegö (KMS) Toeplitz matrix K, whose elements are $k_{ij} = y^{|i-j|}$, $y \in \mathbb{R}$. If $0 < |y| < 1$, the matrix is positive definite. It is ill-conditioned for y close to 1. We choose $y = 0.2$, for which $\kappa(K)$ is around 2.25 when the dimension ranges between 100 and 1000. We also consider the nearly orthogonal Chebyshev Vandermonde-like matrix Q, which is given by the *gallery* function, using the test matrix *orthog* with the parameter $k = -1$. Its elements, based on the extrema of the Chebyshev polynomial T_{n-1}, are $q_{ij} = \cos((i-1)(j-1)\pi/(n-1))$, and we set $A = Q^T Q$. Their condition number lies in the interval $[2.045, 2.147]$ for dimensions between 100 and 1000. Finally, we consider the *parter* matrix P whose elements are $p_{ij} = 1/(i - j + 0.5)$. P is a Cauchy and a Toeplitz matrix. We consider the matrix $A = P^T P$ whose condition number is 10.997 for $p = 100$ and 17.898 for $p = 1000$. In Table 5.12, we give the lower and the upper bounds for $\mathrm{Tr}(A^2)$ for these matrices obtained from Corollary 4.2.

TABLE 5.12
Bounds for $\mathrm{Tr}(A^2)$.

Matrix	Dim.	lower bound	Exact	upper bound
Prolate	100	3.243e2	3.394e2	3.649e2
Prolate	1000	3.240e3	3.399e3	3.645e3
KMS	100	1.002e2	1.083e2	1.175e2
KMS	1000	1.000e3	1.083e3	1.174e3
Q^TQ	100	2.652e5	2.748e5	3.059e5
Q^TQ	1000	2.515e8	2.525e8	2.851e8
P^TP	100	9.446e3	9.544e3	3.091e4
P^TP	1000	9.702e4	9.715e4	4.840e5

6. Concluding remarks. In this paper, we extended the technique presented in [4] for estimating the trace of the inverse of a matrix to the trace of its powers. As explained in Section 1, such estimates have applications in various branches of mathematics. According to the numerical tests we performed, it seems that our estimates are not very sensitive to perturbations on the initial matrix. Probably, they could be further improved by statistical

ETNA
Kent State University
http://etna.math.kent.edu

techniques such as trimmed values or bootstrapping, as already done in [4]. We also performed some tests with $q \in \mathbb{R}$, which proved to be conclusive. The ideas presented in this paper could be extended to a complex Hilbert space. Also, the possible extension of the extrapolation technique for obtaining estimates of $(x, f(A)x)$, where f is some function, has to be studied. Although the numerical results presented here only deal with matrices, they could be extended, under convenient assumptions, to the trace of powers of a positive self-adjoint operator in a Hilbert space.

Acknowledgments. We would like to thank Michele Benzi and Christos Koukouvinos for constructive discussions and advice, and for providing us useful information. We are grateful to the reviewers of this paper whose remarks helped us to extend some results. The second author (P. F.) acknowledges financial support from State Scholarships Foundation (IKY), following a procedure of individualised assessment, funded by the European Social Fund (ESF) and NSRF of 2007–2013.

REFERENCES

[1] H. AVRON, *Counting triangles in large graphs using randomized matrix trace estimation*, in Proceedings of KDD-LDMTA'10, ACM, 2010.

[2] Z. BAI, M. FAHEY, AND G. GOLUB, *Some large-scale matrix computation problems*, J. Comput. Appl. Math., 74 (1996), pp. 71–89.

[3] C. BREZINSKI, *Error estimates for the solution of linear systems*, SIAM J. Sci. Comput., 21 (1999), pp. 764–781.

[4] C. BREZINSKI, P. FIKA, AND 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., to appear, 2012.

[5] C. BREZINSKI AND M. RAYDAN, *Cauchy-Schwartz and Kantorovich type inequalities for scalar and matrix moment sequences*, Adv. Comput. Math., 26 (2007), pp. 71–80.

[6] J.-P. CHEHAB AND M. RAYDAN, *Geometrical properties of the Frobenius condition number for positive definite matrices*, Linear Algebra Appl., 429 (2008), pp. 2089–2097.

[7] M. T. CHU, *Symbolic calculation of the trace of the power of a tridiagonal matrix*, Computing, 35 (1985), pp. 257–268.

[8] B. N. DATTA AND K. DATTA, *An algorithm for computing powers of a Hessenberg matrix and its applications*, Linear Algebra Appl., 14 (1976), pp. 273–284.

[9] T. A. DAVIS AND Y. HU, *The University of Florida Sparse Matrix Collection*, ACM Trans. Math. Software, 38 (2011), 1 (25 pages). Available at http://www.cise.ufl.edu/research/sparse/matrices/.

[10] G. H. GOLUB AND G. MEURANT, *Matrices, Moments and Quadrature with Applications*, Princeton University Press, 2010.

[11] N. HIGHAM, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008.

[12] M. HUTCHINSON, *A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines*, Comm. Statist. Simulation Comput., 18 (1989), pp. 1059–1076.

[13] V. B. LIDSKII, *Non self-adjoint operators with a trace*, Dokl. Akad. Nauk SSSR, 125 (1959), pp. 485–487 (in Russian).

[14] V. PAN, *Estimating the extremal eigenvalues of a symmetric matrix*, Comput. Math. Appl., 20 (1990), pp. 17–22.

[15] F. PUKELSHEIM, *Optimal Design of Experiments*, Wiley, New York, 1993.

[16] B. SIMON, *Trace Ideals and their Applications*, 2nd ed., Amer. Math. Soc., Providence, 2005.

[17] R. TÜRKMEN AND Z. ULUKÖK, *On the Frobenius condition number of positive definite matrices*, J. Inequal. Appl., (2010), 897279 (11 pages).

[18] A. V. ZARELUA, *On congruences for the traces of powers of some matrices*, Proc. Steklov Inst. Math., 263 (2008), pp. 78–98.