Bounds for the Trace of the Inverse and the Determinant of Symmetric Positive Definite Matrices

ZHAOJUN BAI[1, *] AND GENE H. GOLUB[2, †]

[1]*Department of Mathematics,*
University of Kentucky,
Lexington, KY 40506
E-mail: `bai@ms.uky.edu`

[2]*Scientific Computing and Computational Mathematics Program,*
Computer Science Department,
Stanford University,
Stanford, CA 94305
E-mail: `golub@sccm.stanford.edu.`

Dedicated to T. Rivlin on the occasion of his 70th birthday

Lower and upper bounds are given for the trace of the inverse $\mathrm{tr}(A^{-1})$ and the determinant $\det(A)$ of a symmetric positive definite matrix A. They are derived by applying Gaussian quadrature and related theory.

The bounds for $\det(A)$ appears to be new. For the bounds of $\mathrm{tr}(A^{-1})$, the Kantorovich inequality is available for providing such bounds. In a number of examples, our bounds are found to be tighter when simple trial vectors are used in Kantorovich's bound. The new bounds are equivalent to Robinson and Wathen's variational bounds. But our bounds are directly derived for the quantity instead of the summation of bounds for each diagonal entry of A^{-1}.

Subject classification: AMS(MOS) 65F10

Keywords: trace, inverse, determinant, quadrature

1 Introduction

There are a number of applications where it is desired to estimate the bounds for the quantities of the trace of the inverse $\mathrm{tr}(A^{-1})$ and the determinant $\det(A)$ of a matrix A, such as in the study of fractals [14, 18], lattice Quantum Chromodynamics (QCD) [15, 3], crystals [11, 12], the generalized cross-validation and its applications (see [7] and references therein).

In this paper, we focus on deriving lower and upper bounds for the quantities $\mathrm{tr}(A^{-1})$

*The author was supported in part by by an NSF grant ASC-9313958 and in part by an DOE grant DE-FG03-94ER25219.

†The work of this author was in part supported by NSF under grant CCR-9505393.

and $\det(A)$ of a symmetric positive definite matrix A. Throughout this paper, A will denote an n-by-n symmetric positive definite matrix with eigenvalues

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

$\lambda(A)$ is the set of all eigenvalues. The parameters α and β denote the bounds for the smallest and largest eigenvalues λ_1 and λ_n of A,

$$0 < \alpha \leq \lambda_1, \qquad \lambda_n \leq \beta.$$

a_{ij} or $(A)_{ij}$ will denote the (i,j) entry of a matrix A. Using the eigenvalue decomposition and the definition of matrix function [6], it is easy to prove the identity

$$\ln(\det(A)) = \mathrm{tr}(\ln(A)) \tag{1}$$

for a symmetric positive definite matrix A. Therefore, instead of bounding $\det(A)$, we will bound $\ln(\det(A))$, which is turned into bounding $\mathrm{tr}(\ln(A))$. Under this reformulation, the problems of bounding the quantities $\mathrm{tr}(A^{-1})$ and $\det(A)$ are unified to bound $\mathrm{tr}(f(A))$ for $f(\lambda) = \lambda^{-1}$ and $\ln\lambda$, respectively.

We first note that the Kantorovich inequality

$$(x^T A^{-1} x)(x^T A x) \leq \frac{1}{4}\left(\frac{\lambda_1}{\lambda_n} + \frac{\lambda_n}{\lambda_1} + 2\right)(x^T x)^2$$

holds for all vectors x. For derivation of this inequality, see [9]. By using a simple trial vector $x = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ in the inequality and noting that the upper bound is monotonically increasing in λ_n/λ_1, it yields

$$(A^{-1})_{ii} \leq \frac{1}{4a_{ii}}\left(\frac{\alpha}{\beta} + \frac{\beta}{\alpha} + 2\right). \tag{2}$$

The summation of the upper bound for all diagonal entries of A^{-1} gives a *Kantorovich's upper bound* for $\mathrm{tr}(A^{-1})$.

Another approach for bounding $\mathrm{tr}(A^{-1})$ is to use variational functional, Robinson and Wathen [13] show that

$$\frac{1}{\alpha} + \frac{(\alpha - a_{ii})^2}{\alpha(\alpha a_{ii} - s_{ii})} \leq (A^{-1})_{ii} \leq \frac{1}{\beta} - \frac{(a_{ii} - \beta)^2}{\beta(s_{ii} - \beta a_{ii})}, \tag{3}$$

where $s_{ii} = \sum_{k=1}^{n} a_{ik}^2$. Hence the sums of the lower and upper bounds for all diagonal entries of A^{-1} give *Robinson and Wathen's lower and upper bounds* for $\mathrm{tr}(A^{-1})$, respectively. Two other types of bounds for $(A^{-1})_{ii}$ are also presented in [13], but more information is required.

The third approach is to use Gaussian quadrature and related theory. As discussed in [4, 5, 1], one first bounds the quantity $x^T f(A) x$ for a given vector x, then probabilistic bounds and estimates can be obtained for $\mathrm{tr}(f(A))$ by using Monte Carlo simulation. We refer to [1] for details.

The new bounds derived in this paper also use Gaussian quadrature and related theory. But they are exact lower and upper bounds for $\mathrm{tr}(f(A))$ instead of probabilistic bounds as given in [1]. The new bounds are derived by directly considering the quantity $\mathrm{tr}(f(A))$ instead of each diagonal entry of $f(A)$. They are very cheap to compute. In a number of examples our bounds are found to be tighter than Kantorovich's upper bound, and are equivalent to the Robinson and Wathen's bounds, which is computationally more expensive. Our experiences indicate that the probabilistic bounds presented in [1] are the most accurate, but they are also the most expensive ones in terms of computational costs and memory requirement.

In section 2 we present the lower and upper bounds for $\mathrm{tr}(A^{-1})$ and $\mathrm{tr}(\ln(A))$. A number of examples, coming from the different applications, and comparisons with the Kantorovich's upper bound, Robinson and Wathen's bounds and the estimates using Monte Carlo simulation are given in section 3. We give concluding remarks in section 4.

2 Bounds for tr(A^{-1}) and tr(ln(A))

Let

$$\mu_r = \mathrm{tr}(A^r) = \sum_{i=1}^{n} \lambda_i^r = \int_{\alpha}^{\beta} \lambda^r \, d\gamma(\lambda), \tag{4}$$

where the weight function $\gamma(\lambda)$ of the Stieltjes integral is $\gamma(\lambda) = \sum_{j=1}^{n} I(\lambda - \lambda_j)$, and $I(\lambda)$ is the unit step function: $I(\lambda) = 0$ if $\lambda < 0$ and $I(\lambda) = 1$ if $\lambda \geq 0$. Note that one can easily compute

$$\mu_0 = n, \qquad \mu_1 = \sum_{i=1}^{n} a_{ii}, \qquad \mu_2 = \sum_{i,j=1}^{n} a_{ij}^2 = \|A\|_F^2.$$

Our first task is to use μ_0, μ_1 and μ_2 and the parameters α and β to determine a lower and an upper bound for $\mu_{-1} = \mathrm{tr}(A^{-1})$.

The approach is to use the classical Gaussian quadrature and related theory, see for example [2]. Specifically, we use the Gauss-Radau quadrature rule. By the rule, the integral in (4) can be written as

$$\mu_r = \int_{\alpha}^{\beta} \lambda^r \, d\gamma(\lambda) = \bar{\mu}_r + \mathcal{R}[\mu_r], \tag{5}$$

where $\bar{\mu}_r$ is the following quadrature formula

$$\bar{\mu}_r = w_0 t_0^r + w_1 t_1^r, \tag{6}$$

w_0 and w_1 are weights and to be determined. t_0 and t_1 are nodes. The node t_0 is prescribed, say $t_0 = \alpha$ or β. t_1 is unknown and to be determined. The remainder

$$\mathcal{R}[\mu_r] = \frac{1}{6} r(r-1)(r-2)\eta^{r-3} \int_{\alpha}^{\beta} (\lambda - t_0)(\lambda - t_1)^2 d\gamma(\lambda)$$

for some $\alpha < \eta < \beta$. If $\mathcal{R}[\mu_r] \geq 0$, $\bar{\mu}_r$ is a lower bound of μ_r and if $\mathcal{R}[\mu_r] \leq 0$, $\bar{\mu}_r$ is a upper bound of μ_r.

From (6), we see that $\bar{\mu}_r$ satisfies a second order difference equation

$$c\bar{\mu}_r + d\bar{\mu}_{r-1} - \bar{\mu}_{r-2} = 0 \tag{7}$$

for certain coefficients c and d. The nodes t_0 and t_1 are the roots of the characteristic polynomial

$$p(\xi) = c\xi^2 + d\xi - 1. \tag{8}$$

To determine the coefficients c and d, by using (7) with the fact $\bar{\mu}_r = \mu_r$ for $r = 0, 1, 2$, and the prescribed node t_0 being the root of the characteristic polynomial (8), we have

$$\begin{aligned} c\mu_2 + d\mu_1 - \mu_0 &= 0, \\ ct_0^2 + dt_0 - 1 &= 0. \end{aligned}$$

Solving the above linear equations for c and d yields

$$\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \mu_2 & \mu_1 \\ t_0^2 & t_0 \end{bmatrix}^{-1} \begin{bmatrix} \mu_0 \\ 1 \end{bmatrix}$$

Once having the coefficients c and d, the node t_1 of the quadrature $\bar{\mu}_r$ is given by $t_1 = -1/(t_0 c)$.

For determining the weights w_0 and w_1, we note that

$$\begin{aligned} \mu_1 &= w_0 t_0 + w_1 t_1, \\ \mu_2 &= w_0 t_0^2 + w_1 t_1^2. \end{aligned}$$

Then

$$\begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{bmatrix} t_0 & t_1 \\ t_0^2 & t_1^2 \end{bmatrix}^{-1} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}.$$

To bound $\mathrm{tr}(A^{-1}) = \mu_{-1}$, writing the difference equation (7) with $r = 1$, we have

$$c\bar{\mu}_1 + d\bar{\mu}_0 - \bar{\mu}_{-1} = 0.$$

i.e.,

$$\bar{\mu}_{-1} = c\bar{\mu}_1 + d\bar{\mu}_0 = \begin{bmatrix} \mu_1 & \mu_0 \end{bmatrix} \begin{bmatrix} \mu_2 & \mu_1 \\ t_0^2 & t_0 \end{bmatrix}^{-1} \begin{bmatrix} \mu_0 \\ 1 \end{bmatrix}.$$

By the Gauss-Radau quadrature rule (5) with $r = -1$, we have

$$\mu_{-1} = \bar{\mu}_{-1} + \mathcal{R}[\lambda^{-1}],$$

where the remainder

$$\mathcal{R}[\lambda^{-1}] = -\frac{1}{\eta^4} \int_\alpha^\beta (\lambda - t_0)(\lambda - t_1)^2 d\gamma(\lambda)$$

for some $\alpha < \eta < \beta$. If the prescribed node $t_0 = \alpha$, $\mathcal{R}[\lambda^{-1}] \le 0$, then $\bar{\mu}_{-1}$ is an upper bound of μ_{-1}. If $t_0 = \beta$, $\mathcal{R}[\lambda^{-1}] \ge 0$, then $\bar{\mu}_{-1}$ is a lower bound of μ_{-1}. In summary, we have the following bounds for $\mathrm{tr}(A^{-1})$.

Theorem 1 **(Lower and upper bounds for tr(A^{-1}))**
Let A be an n-by-n symmetric positive definite matrix, $\mu_1 = tr(A)$, $\mu_2 = \|A\|_F^2$ and $\lambda(A) \subset [\alpha, \beta]$ with $\alpha > 0$, then

$$\begin{bmatrix} \mu_1 & n \end{bmatrix} \begin{bmatrix} \mu_2 & \mu_1 \\ \beta^2 & \beta \end{bmatrix}^{-1} \begin{bmatrix} n \\ 1 \end{bmatrix} \leq tr(A^{-1}) \leq \begin{bmatrix} \mu_1 & n \end{bmatrix} \begin{bmatrix} \mu_2 & \mu_1 \\ \alpha^2 & \alpha \end{bmatrix}^{-1} \begin{bmatrix} n \\ 1 \end{bmatrix} \tag{9}$$

Let us turn to our second task for bounding $\mathrm{tr}(\ln(A))$. Note that the identity (1) can be further written as

$$\ln(\det(A)) = \mathrm{tr}(\ln A) = \sum_{i=1}^{n} (\ln \lambda_i) = \int_\alpha^\beta (\ln \lambda) d\gamma(\lambda),$$

where the weight function $\gamma(\lambda)$ of the Stieltjes integral is the same as the weight function defined in (4). Again, by using the Gauss-Radau quadrature rule, we have

$$\mathrm{tr}(\ln(A)) = \int_\alpha^\beta (\ln \lambda) d\gamma(\lambda) = \mathcal{I}[\ln \lambda] + \mathcal{R}[\ln \lambda],$$

where the quadrature term $\mathcal{I}[\ln \lambda]$ is

$$\mathcal{I}[\ln \lambda] = w_0 \ln(t_0) + w_1 \ln(t_1).$$

The remainder

$$\mathcal{R}[\ln \lambda] = \frac{2}{3} \frac{1}{\eta^3} \int_\alpha^\beta (\lambda - t_0)(\lambda - t_1)^2 d\gamma(\lambda)$$

for some $\alpha < \eta < \beta$. Therefore, if $t_0 = \alpha$, $\mathcal{R}[\ln \lambda] \geq 0$, then $\mathcal{I}[\ln \lambda]$ is a lower bound of $\mathrm{tr}(\ln(A))$. If $t_0 = \beta$, $\mathcal{R}[\ln \lambda] \leq 0$, then $\mathcal{I}[\ln \lambda]$ is an upper bound. Therefore, we derive the following bounds for $\mathrm{tr}(\ln(A))$.

Theorem 2 **(Lower and upper bounds for tr(ln(A)))**
Let A be an n-by-n symmetric positive definite matrix, $\mu_1 = tr(A)$, $\mu_2 = \|A\|_F^2$ and $\lambda(A) \subset [\alpha, \beta]$ with $\alpha > 0$, then

$$\begin{bmatrix} \ln \alpha & \ln \underline{t} \end{bmatrix} \begin{bmatrix} \alpha & \underline{t} \\ \alpha^2 & \underline{t}^2 \end{bmatrix}^{-1} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \leq tr(\ln(A)) \leq \begin{bmatrix} \ln \beta & \ln \bar{t} \end{bmatrix} \begin{bmatrix} \beta & \bar{t} \\ \beta^2 & \bar{t}^2 \end{bmatrix}^{-1} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \tag{10}$$

where

$$\underline{t} = \frac{\alpha \mu_1 - \mu_2}{\alpha n - \mu_1} \qquad \text{and} \qquad \bar{t} = \frac{\beta \mu_1 - \mu_2}{\beta n - \mu_1}$$

The bounds (9) and (10) involve only the trace of the lower orders of the matrix power A^r, namely, $\mu_0 = \mathrm{tr}(A^0) = n$, $\mu_1 = \mathrm{tr}(A^1)$ and $\mu_2 = \mathrm{tr}(A^2)$. They can be easily computed. If the trace of the higher orders of the matrix power A^r, $r \geq 3$, are available, then using the principles discussed above, we can derive tighter bounds.

The parameters α and β for the bounds of eigenvalues of A must be provided, which happens to be required in all such bounds discussed in Section 1.

Table 1
Lower and upper bounds for tr(A^{-1})

Matrix (order)	"Exact"	MC estimation	Lower bound	Upper bound
Poisson (900)	$5.12644 \cdot 10^2$	$5.02012 \cdot 10^2$	$2.60852 \cdot 10^2$	$8.74445 \cdot 10^3$
Wathen (341)	$6.16011 \cdot 10^2$	$6.21092 \cdot 10^2$	$4.49424 \cdot 10^2$	$9.29451 \cdot 10^2$
Heat flow (625)	$3.65722 \cdot 10^2$	$3.65179 \cdot 10^2$	$3.59979 \cdot 10^2$	$3.73996 \cdot 10^2$

3 Examples

In this section, we use four examples to show the tightness of the bounds given in (9) and (10). We will also compare with the Kantorovich's upper bound (2), Robinson and Wathen's bounds (3) and the approach using Monte Carlo simulation (henceforth the MC estimation) described in [1].

Numerical experiments are carried out in Matlab environment on a SUN Sparcstation 10. The so-called "exact" value of tr(A^{-1}) is computed by analytic formulas if available, or by first computing the inverse using function `inv` in Matlab, and then calculating the trace. For the "exact" value of $\mathrm{tr}(\ln(A)) = \ln(\det(A))$, we use the analytic formula if available, or first compute the Cholesky decomposition of A using Matlab function `chol` and then compute the natural logarithm of the product of diagonals of Cholesky factor.

Example 1 (Pei matrix): Consider the so-called n-by-n Pei matrix $A = \tau I + uu^T$, where $u = (1, 1, \ldots, 1)^T$ [8]. It is easy to see that A has two distinct eigenvalues τ and $n + \tau$. The eigenvalue τ has multiplicity $n - 1$. If $\tau > 0$, A is symmetric positive definite. By the Sherman-Morrison formula [6], the inverse of A can be written as

$$A^{-1} = \frac{1}{\tau} I - \frac{1}{\tau(\tau + n)} uu^T$$

Then $\mathrm{tr}(A^{-1}) = \frac{n}{\tau} - \frac{n}{\tau(\tau+n)}$ and $\mathrm{tr}(\ln(A)) = (n-1)\ln\tau + \ln(\tau+n)$. It is easy to compute that $\mu_1 = \mathrm{tr}(A) = (\tau+1)n$ and $\mu_2 = \mathrm{tr}(A^2) = n^2 + \tau(\tau+2)n$. If let parameters $\alpha = \tau$ and $\beta = n + \tau$, then by straightforward algebraic calculation, both lower and upper bounds for tr(A^{-1}) in (9) are equal to the exact value. Similarly, one can also easily show that both lower and upper bounds for tr(ln(A)) in (10) are also equal to the exact value. For this example, the bounds (9) and (10) are perfect! Of course, one can also verify that for this example, there are no intergration errors (the remainders are zero) in the Gauss-Radau quadrature rule.

For this example, Robinson and Wathen's bounds for tr(A^{-1}) are also equal to the exact value.

Example 2 (Poisson matrix): The matrix of order m^2 is a block tridiagonal matrix from the 5-point central difference discretization of the 2-D Poisson's equation on a $m \times m$ square mesh [8]. It can be shown that the parameters α and β for the bounds of the smallest and largest eigenvalues are $\alpha = 2\left(\frac{\pi}{m+1}\right)^2$ and $\beta = 8$, respectively. In Tables 1 and 2, we have tabulated the exact values for tr(A^{-1}) and tr(ln(A)), the estimated values by the MC estimation [1]. and the lower and upper bounds given in (9) and (10) for the 900 by 900 Poisson matrix (i.e., $m = 30$).

Table 2
Lower and upper bounds for $\mathrm{tr}(\ln(A))$

Matrix (order)	"Exact"	MC estimation	Lower bound	Upper bound
Poisson (900)	$1.06500 \cdot 10^3$	$1.06023 \cdot 10^3$	$4.73862 \cdot 10^2$	$1.16857 \cdot 10^3$
Wathen (341)	$-1.20071 \cdot 10^2$	$-1.20263 \cdot 10^2$	$-2.00165 \cdot 10^2$	$-6.86595 \cdot 10^1$
Heat flow (625)	$3.51679 \cdot 10^2$	$3.50715 \cdot 10^2$	$3.47348 \cdot 10^2$	$3.54997 \cdot 10^2$

The Kantorovich's upper bound for $\mathrm{tr}(A^{-1})$ is $2.20208 \cdot 10^4$. Robinson and Wathen's lower and upper bounds are $2.60969 \cdot 10^2$ and $8.73279 \cdot 10^3$. Note that the estimates from the Monte Carlo simulation for $\mathrm{tr}(A^{-1})$ and $\mathrm{tr}(\ln(A))$ are only at 2% and 0.4% of relative errors of the actual values, respectively [1].

Example 3 (Wathen matrix): The matrix A is "$\mathtt{wathen}(n_x, n_y)$" in the set of Matlab test matrices collection by Higham [8]. It is a consistent mass matrix in finite element computations for a regular n_x-by-n_y grid of 8-node (serendipity) elements in 2 space dimensions (see [16]). The resulting matrix is of order $n = 3n_xn_y + 2n_x + 2n_y + 1$. Let D be the diagonals of A, then realistic bounds for the eigenvalues of $D^{-1}A$ are given by Wathen [17]. In our numerical experiment, we let $n_x = n_y = 10$. Then the matrix A is of order $n = 341$ and $\alpha = 0.25$ and $\beta = 4.5$. The bounds for $\mathrm{tr}(D^{\frac{1}{2}}A^{-1}D^{\frac{1}{2}})$ and $\mathrm{tr}(\ln(D^{-\frac{1}{2}}AD^{-\frac{1}{2}}))$ are tabulated in Table 1 and 2.

The Kantorovich's upper bound for $\mathrm{tr}(D^{\frac{1}{2}}A^{-1}D^{\frac{1}{2}})$ is $1.70974 \cdot 10^3$ and the Robinson and Wathen's lower and upper bounds are $4.53280 \cdot 10^2$ and $9.19952 \cdot 10^2$, respectively. Again, note that the estimated values of the MC estimation are only at 0.8% and 0.1% of relative errors of the actual values.

Example 4: This test matrix is from [10], see also [13]. The matrix is resulted from the implicit finite different discretization of a linear heat flow problem. It is a m^2 by m^2 block tridiagonal matrix of the form

$$A = \begin{pmatrix} D & C & & \\ C & D & \ddots & \\ & \ddots & \ddots & C \\ & & C & D \end{pmatrix},$$

where D is a $m \times m$ tridiagonal matrix with $1 + 4\nu$ on diagonal, and $-\nu$ on super- and sub-diagonal, and C is a diagonal matrix with diagonal entries ν. ν is the ratio of the time step and the square of grid size. The Gershgorin circle theorem gives $\alpha = 1$ and $\beta = 1 + 8\nu$ for the eigenvalue bounds of A. We have tabulated in Tables 1 and 2 the bounds for $\mathrm{tr}(A^{-1})$ and $\mathrm{tr}(\ln(A))$, respectively, where $n = 625$ ($m = 25$) and $\nu = 0.5$.

The Kantorovich's upper bound for $\mathrm{tr}(A^{-1})$ is $4.32692 \cdot 10^2$ and the Robinson and Wathen's lower and upper bounds are $3.59996 \cdot 10^2$ and $3.73972 \cdot 10^2$, respectively.

4 Concluding Remarks

Simple bounds for $\mathrm{tr}(A^{-1})$ and $\mathrm{tr}(\ln(A))$ of a symmetric positive definite matrix A are derived by using Gaussian quadrature and related theory. The bounds involve only n (the order of A), $\mathrm{tr}(A)$ and $||A||_F^2$ and the parameters α and β, namely for the bounds of eigenvalues of A.

From the numerical examples presented in Section 3 and numerous other experiments conducted, our bounds for $\mathrm{tr}(A^{-1})$ are found to be tighter when simple trial vectors are used in the Kantorovich's bound, and are equivalent to the Robinson and Wathen's bounds. But it is generally poorer than the probablistic bounds and estimations for $\mathrm{tr}(A^{-1})$ and $\ln(\det(A))$ derived by using Gaussian quadrature and Monte Carlo simulation [1]. The latter costs significantly more arithmetic operations and memory. But a fully parallelism scheme can be developed for the simulation [1].

We point out that using Gaussian quadrature and related theory, we have the advantage of easily extending the approach for $\mathrm{tr}(A^{-1})$ to $\mathrm{tr}(\ln(A))$, while the approaches based on Kantorovich inequality and variational inequality do not enjoy. If the traces of higher orders of the matrix power A^r, $r \geq 3$, are available, then bounds can be further tightened by using the same technique. Moreover, one could use modified moments to get improved estimates of the quadrature rule.

5 Afternote

While we were finishing up this paper, we read an article by Ortner and Kräuter on lower bounds for the determinant and the trace of the inverse in the most recent issue of *Linear Algebra and its Applications* [12]. One central problem studied by Ortner and Kräuter is to find lower bounds of $\mathrm{tr}(P^{-1})$ and $\det(P^{-1})$, where $P = \frac{1}{m}X^TX$, X is a given m-by-n ($m \geq n$) full rank matrix, whose rows have unit length. This problem arises from accuracy considerations in real second-rank tensor measurements of single crystals [11]. By using standard matrix theory, one can show that the condition number of the matrix X is related to the quantity $\mathrm{tr}(P^{-1})$, namely,

$$\kappa_F(X) = ||X||_F||X^\dagger||_F = \sqrt{\mathrm{tr}(P^{-1})}.$$

where $X^\dagger$ is the Moore-Penrose inverse of X [6]. Therefore, a lower bound of $\mathrm{tr}(P^{-1})$ also gives a lower bound of the condition number of X.

Using an approach based on matrix theory and combinatorics, various lower bounds of $\mathrm{tr}(P^{-1})$ are derived in [12]. The sharpness of those lower bounds are demonstrated for small m and n ([12, Example 3]). Our approach yields the same lower bounds for their set of test problems, provided that the extreme eigenvalues of P are available. As indicated by Ortner and Kräuter [12], in most cases, it is very hard, if not impossible, to find a suitable configuration of the row vectors of X to attain the optimal lower bound. Since our approach gives both lower and upper bounds of the quantity $\mathrm{tr}(P^{-1})$, it might provide a way to assess how far a given configuration is from the optimal configuration. It would be interesting to make further investigations in this direction.

References

[1] Z. Bai, M. Fahey, and G. Golub. Some large scale matrix computation problems. *J. of Comput. and Appl. Math.*, 1996. to appear.

[2] P. Davis and P. Rabinowitz. *Methods of Numerical Integration*. Academic Press, NY, 1984.

[3] S. Dong and K. Liu. Stochastic estimation with z_2 noise. *Physics Letters B*, 328 (1994) 130–136.

[4] G. Golub and G. Meurant. Matrices, moments and quadrature. in *Proceedings of the 15th Dundee Conference, June 1993*, D. F. Griffiths and G. A. Watson, Eds., Longman Scientific & Technical, 1994.

[5] G. Golub and Z. Strakos. Estimates in quadratic formulas. *Numerical Algorithms*, 8 (1994) 241 - 268.

[6] G. Golub and C. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, 2nd edition, 1989.

[7] G. Golub and U. Von Matt. Generalized cross-validation for large scale problems. Report SCCM-96-05, Stanford University, Stanford, 1996.

[8] N. J. Higham. The test matrix toolbox for Matlab. Numerical Analysis Report 237, University of Manchester, Manchester, 1993.

[9] A. S. Householder. *The Theory of Matrices in Numerical Analysis*. Blaisdell, New York, NY, 1964. Dover edition 1975.

[10] A. R. Mitchell and D. F. Griffiths. *The Finite Difference Method in Partial Differential Equations*. Wiley, NY, 1980.

[11] B. Ortner. On the selection of measurement directions in second-rank tensor (e.g. elastic strain) determination of single crystals. *J. Appl. Cryst.*, 22 (1989) 216–221.

[12] B. Ortner and A. R. Kräuter. Lower bounds for the determinant and the trace of a class of Hermitian matrices. *Lin. Alg. and Appl.*, 236 (1996) 147–180.

[13] P. Robinson and A. J. Wathen. Variational bounds on the entries of the inverse of a matrix. *IMA Journal of Numerical Analysis*, 12 (1992) 463–486.

[14] B. Sapoval, Th. Gobron, and A. Margolina. Vibrations of fractal drums. *Phy. Rev. Lett.*, 67 (1991) 2974–2977.

[15] J. C. Sexton and D. H. Weingarten. Systematic expansion for full QCD based on the valence approximation, IBM T. J. Watson Research Center, Preprint, 1994

[16] G. Strang and G. Fix. *An Analysis of the Finite Element Method*. Prentice Hall, Englewood Cliffs, NJ, 1973.

[17] A. J. Wathen. Realistic eigenvalue bounds for the Galerkin mass matrix. *IMA Journal of Numerical Analysis*, 7 (1987) 449–457.

[18] S. Y. Wu, J. A. Cocks, and C. S. Jayanthi. An accelerated inversion algorithm using the resolvent matrix method. *Comput. Phy. Comm.*, 71 (1992) 125–133.