ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 429 (2008) 519–526
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

On lower bounds for the largest eigenvalue of a symmetric matrix

S.G. Walker [a,*], P. Van Mieghem [b]

[a] *Institute of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury CT2 7NZ, UK*
[b] *Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science, P.O. Box 5031, 2600 GA Delft, The Netherlands*

Received 17 September 2007; accepted 13 March 2008
Available online 22 April 2008
Submitted by S. Friedland

Abstract

We consider lower bounds for the largest eigenvalue of a symmetric matrix. In particular we extend a recent approach by Piet Van Mieghem.

AMS classification: Primary 15A42; Secondary 30B10

Keywords: Largest eigenvalue; Positive definite matrix; Symmetric matrix

1. Introduction

Let $\lambda_{\max}(A)$ be the largest eigenvalue of a symmetric $m \times m$ matrix $A = (a_{ij})$. Since

$$\lambda_{\max}(A) = \max_{x \neq 0} \frac{x^{\mathrm{T}} A x}{x^{\mathrm{T}} x}$$

it clearly follows that a lower bound for $\lambda_{\max}(A)$ is given by

$$\lambda_{\max}(A) \geqslant \frac{u^{\mathrm{T}} A u}{u^{\mathrm{T}} u} \tag{1}$$

* Corresponding author. Tel.: +44 (0) 1227 823800; fax: +44 (0) 1227 827932.
E-mail addresses: S.G.Walker@kent.ac.uk (S.G. Walker), P.F.A.VanMieghem@tudelft.nl (P. Van Mieghem).

0024-3795/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2008.03.007

where $u^{\mathrm{T}} = (1 \cdots 1)$. Note that

$$N_1 = u^{\mathrm{T}} A u = \sum_{ij} a_{ij},$$

$u^{\mathrm{T}} u = m$ and N_1/m is a commonly used lower bound for $\lambda_{\max}(A)$. Recent work on lower bounds for a symmetric matrix has been done by Van Mieghem [2]. He showed that

$$\lambda_{\max}(A) \geqslant \frac{N_1}{m} + 2\left(\frac{N_3}{2m} - \frac{N_1 N_2}{m^2} + \frac{N_1^3}{2m^3}\right)\lambda_0^{-2} + \mathrm{O}(t^{-4}), \tag{2}$$

where $t \geqslant T$, $\lambda_0 = t\sqrt{m}$,

$$T = \frac{1}{\sqrt{m}} \max_{1 \leqslant j \leqslant m} \left(a_{jj} + \sum_{i \neq j} |a_{ij}|\right), \tag{3}$$

and $N_k = u^{\mathrm{T}} A^k u$ with $N_0 = m$.

The aim in the current paper is to extend the results of Van Mieghem [2]. The central idea of the paper is to apply the classic bound to transforms of A. Applying standard bounds to transformed matrices which result in improved bounds has recently been exploited in Walker [3,4] and Liu et al. [1]. We derive the general lower bound in Section 2, where we also consider some specific cases. Section 3 provides a further useful result when A is positive definite and finally Section 4 concludes with a numerical example.

2. Lower bounds for symmetric matrices

Consider the $m \times m$ symmetric matrix

$$A_t = \sum_{k=0}^{\infty} f_k A^k t^{-k},$$

where the Taylor series $\sum_{k=0}^{\infty} f_k x^k = f(x)$ converges for $|x| < R_f$, where $R_f > 0$ is the radius of convergence. If λ is an eigenvalue of A, corresponding to eigenvector v, then

$$A_t v = \sum_{k=0}^{\infty} f_k A^k t^{-k} v = \sum_{k=0}^{\infty} f_k \lambda^k t^{-k} v = f\left(\frac{\lambda}{t}\right) v.$$

The series converges for any eigenvalue of A provided we choose $t > \tilde{\lambda}/R_f$, where $\tilde{\lambda} = \max_{1 \leqslant j \leqslant m}\{|\lambda_j|\}$.

If $f(x)$ is real for real x and increasing, then $\lambda_{\max}(A_t) = f\left(\frac{\lambda_{\max}(A)}{t}\right)$. Next, we apply the classical bound (1) to A_t and obtain

$$\lambda_{\max}(A_t) \geqslant \frac{u^{\mathrm{T}} A_t u}{m} = \frac{1}{m} \sum_{k=0}^{\infty} f_k (u^{\mathrm{T}} A^k u) t^{-k} = \frac{1}{m} \sum_{k=0}^{\infty} f_k N_k t^{-k}.$$

It follows from (1) that $N_k \leqslant m\lambda_{\max}(A^k)$. Since $\lambda_{\max}(A^k) \leqslant \tilde{\lambda}^k$, we have that $N_k \leqslant m\tilde{\lambda}^k$ and this inequality shows that the series $\sum_{k=0}^{\infty} f_k N_k t^{-k}$ indeed converges for $t > \tilde{\lambda}/R_f$.

Since also the inverse function $f^{-1}(x)$ is increasing when $f(x)$ is increasing such that

$$\lambda_{\max}(A) = t f^{-1}(\lambda_{\max}(A_t)),$$

we arrive at the inequality

$$\lambda_{\max}(A) \geqslant t f^{-1}\left(\frac{1}{m}\sum_{k=0}^{\infty} f_k N_k t^{-k}\right). \tag{4}$$

The best possible bound is reached when the right hand side in (4) is optimized over all increasing functions f. Obviously the set of increasing functions includes the case $f(x) = x$ and for this increasing function we obtain the classic inequality $\lambda_{\max}(A) \geqslant N_1/m$. Hence (4) is at least as good as the classic bound when optimized over all increasing functions. In fact as we will see in Section 3, when A is positive definite, it turns out that the worst f is indeed $f(x) = x$.

The function $f^{-1}\left(\frac{1}{m}\sum_{k=0}^{\infty} f_k N_k z^k\right)$ is expanded in a series around $z = 1/t = 0$ in Appendix A to obtain

$$\lambda_{\max}(A) \geqslant t f^{-1}\left(\frac{1}{m}\sum_{k=0}^{\infty} f_k N_k t^{-k}\right) = \sum_{k=1}^{\infty} c_k t^{1-k} \tag{5}$$

and the general term c_k is given in (8). Explicitly, the first few coefficients c_k are

$$\begin{aligned}
c_1 &= \frac{N_1}{m}\\
c_2 &= \frac{f_2}{f_1}\left(\frac{N_2}{m} - \frac{N_1^2}{m^2}\right)\\
c_3 &= \frac{f_3}{f_1}\left(\frac{N_3}{m} - \frac{N_1^3}{m^3}\right) + \frac{2f_2^2}{f_1^2}\left(\frac{N_1^3}{m^3} - \frac{N_1 N_2}{m^2}\right)\\
c_4 &= \frac{f_4}{f_1}\left(\frac{N_4}{m} - \frac{N_1^4}{m^4}\right) + \frac{f_2 f_3}{f_1^2}\left(\frac{5N_1^4}{m^4} - \frac{3N_1^2 N_2}{m^3} - \frac{2N_1 N_3}{m^2}\right)\\
&\quad + \frac{f_2^3}{f_1^3}\left(-\frac{5N_1^4}{m^4} + \frac{6N_1^2 N_2}{m^3} - \frac{N_2^2}{m^2}\right).
\end{aligned}$$

If $R_{f^{-1}}$ is the radius of convergence of the Taylor series of $f^{-1}(x)$ around f_0, then

$$f^{-1}\left(\frac{1}{m}\sum_{k=0}^{\infty} f_k N_k t^{-k}\right) = f^{-1}\left(f_0 + \frac{1}{m}\sum_{k=1}^{\infty} f_k N_k t^{-k}\right)$$

indicates that convergence requires that $\frac{1}{m}\sum_{k=1}^{\infty} f_k N_k t^{-k} < R_{f^{-1}}$. Using $N_k \leqslant m\tilde{\lambda}^k$, the series is bounded by

$$\frac{1}{m}\sum_{k=1}^{\infty} f_k N_k t^{-k} \leqslant \sum_{k=1}^{\infty} f_k \tilde{\lambda}^k t^{-k} = f\left(\frac{\tilde{\lambda}}{t}\right) - f_0$$

from which $f(\frac{\tilde{\lambda}}{t}) < f_0 + R_{f^{-1}}$ and thus, that $t > \frac{\tilde{\lambda}}{f^{-1}(f_0 + R_{f^{-1}})}$. Combined with the above bounds on t, convergence of $\sum_{k=1}^{\infty} c_k t^{1-k}$ requires that

$$t > \tilde{\lambda}\max\left(\frac{1}{R_f}, \frac{1}{f^{-1}\left(f_0 + R_{f^{-1}}\right)}\right) \tag{6}$$

and, in practice, $t > \widetilde{T}\sqrt{m}\max\left(\frac{1}{R_f}, \frac{1}{f^{-1}(f_0+R_{f-1})}\right)$, where

$$\widetilde{T}\sqrt{m} = \max_{1\leqslant j\leqslant m}\left\{\sum_{i=1}^{m}|a_{ij}|\right\},$$

since it is well known that $\tilde{\lambda} < \widetilde{T}\sqrt{m}$.

2.1. Examples

If $f_k = 1$, then $f(x) = \frac{1}{1-x}$ and $f^{-1}(x) = 1 - \frac{1}{x}$. The Taylor series of $f(x)$ around $x = 0$ has $R_f = 1$, while the Taylor series of $f^{-1}(x)$ around $f(0) = 1$ has radius of convergence $R_{f^{-1}} = 1$. Hence, the bound (6) for t yields $t > 2\widetilde{T}\sqrt{m}$ and we find from (5)

$$\lambda_{\max}(A) \geqslant \frac{N_1}{m} + \frac{1}{t}\left(\frac{N_2}{m} - \frac{N_1^2}{m^2}\right) + 2\left(\frac{N_3}{2m} - \frac{N_1N_2}{m^2} + \frac{N_1^3}{2m^3}\right)\frac{1}{t^2} + \mathrm{O}(t^{-3}). \tag{7}$$

The bound (7) is very similar to the Van Mieghem [2] expression (2), except we have an additional $1/t$ term which is positive. Note that the $1/t^2$ term is not necessarily positive. On the other hand, the bound on t in (2) is less than half as large as $2\widetilde{T}\sqrt{m}$ here.

If we choose $f(x) = (1-x)^\alpha$, then the Taylor coefficients around $x = 0$ are $f_k = (-1)^k\binom{\alpha}{k}$ and $R_f = 1$. The inverse function $f^{-1}(x) = 1 - x^{\frac{1}{\alpha}}$ has a radius of convergence around $f(0) = 1$ equal to $R_{f^{-1}} = 1$. Using (6), we have that $t > \widetilde{T}\sqrt{m}\max\left(1, \frac{1}{1-2^{\frac{1}{\alpha}}}\right)$. For $\alpha = -|\beta| < 0$, where $f_k = \binom{|\beta|-1+k}{k}$ and $t > \frac{\widetilde{T}\sqrt{m}}{1-2^{-\frac{1}{|\beta|}}}$, the lower bound (5) up to $\mathrm{O}(t^{-3})$ is

$$\begin{aligned}\lambda_{\max}(A) \geqslant{}& \frac{N_1}{m} + \left(\frac{N_2}{m} - \frac{N_1^2}{m^2}\right)\frac{(|\beta|+1)}{2t}\\ &+ \left\{\frac{(|\beta|+2)}{3(|\beta|+1)}\left(\frac{N_3}{m} - \frac{N_1^3}{m^3}\right) + \left(\frac{N_1^3}{m^3} - \frac{N_1N_2}{m^2}\right)\right\}\frac{(|\beta|+1)^2}{2t^2}.\end{aligned}$$

To compare with (7) where $|\beta| = 1$, we write $t = t_1\frac{1}{2(1-2^{-\frac{1}{|\beta|}})}$, where $t_1 > 2\widetilde{T}\sqrt{m}$,

$$\begin{aligned}\lambda_{\max}(A) \geqslant{}& \frac{N_1}{m} + \left(\frac{N_2}{m} - \frac{N_1^2}{m^2}\right)\frac{(|\beta|+1)\left(1-2^{-\frac{1}{|\beta|}}\right)}{t_1}\\ &+ \left\{\frac{(|\beta|+2)}{3(|\beta|+1)}\left(\frac{N_3}{m} - \frac{N_1^3}{m^3}\right) + \left(\frac{N_1^3}{m^3} - \frac{N_1N_2}{m^2}\right)\right\}\frac{2(|\beta|+1)^2\left(1-2^{-\frac{1}{|\beta|}}\right)^2}{t_1^2}\end{aligned}$$

This shows that the coefficient of $\frac{1}{t_1}$ is larger than in the $\beta = 1$ case provided $|\beta| < 1$. In that case, however, the coefficient of $\frac{1}{t_1^2}$ has a smaller positive $\frac{(|\beta|+2)}{3(|\beta|+1)}$ factor. The argument shows that, depending on the values of N_k, we may fine-tune β to produce a larger lower bound.

Finally, consider $f(x) = e^{ax}$ for which $f_k = \frac{a^k}{k!}$ and $R_f \to \infty$. The inverse function $f^{-1}(x) = \frac{1}{a}\log x$ has a Taylor series around $f(0) = 1$ with $R_{f^{-1}} = 1$. The bound (6) becomes $t > \frac{a\widetilde{T}\sqrt{m}}{\log 2}$ and (5) up to $\mathrm{O}(t^{-3})$ is

$$\lambda_{\max}(A) \geqslant \frac{N_1}{m} + \frac{1}{2}\left(\frac{N_2}{m} - \frac{N_1^2}{m^2}\right)\frac{a}{t} + \left\{\frac{a}{3}\left(\frac{N_3}{m} - \frac{N_1^3}{m^3}\right) + \left(\frac{N_1^3}{m^3} - \frac{N_1 N_2}{m^2}\right)\right\}\frac{a^2}{2t^2}$$

Comparison with (7) via $t = \frac{at_1}{2\log 2}$ gives

$$\lambda_{\max}(A) \geqslant \frac{N_1}{m} + \left(\frac{N_2}{m} - \frac{N_1^2}{m^2}\right)\frac{\log 2}{t_1} + \left\{\frac{a}{3}\left(\frac{N_3}{m} - \frac{N_1^3}{m^3}\right) + \left(\frac{N_1^3}{m^3} - \frac{N_1 N_2}{m^2}\right)\right\}\frac{2\log^2 2}{t_1^2}.$$

The coefficient of $\frac{1}{t_1}$ is now smaller than in the $\beta = 1$ case, but the value of a can be freely chosen in the coefficient of $\frac{1}{t_1^2}$ (ignoring higher order terms).

Another possible sequence of functions to consider is $f(x) = x^k$ for odd k. As has been mentioned the case $k = 1$ provides the classic bound. For these functions the inverse is trivial and hence bounds are easily available.

3. Positive definite case

When A is positive definite we have the following key result:

Lemma 3.1. *It is that $N_k \geqslant N_1^k/m^{k-1}$ for all $k = 1, 2, \ldots$*

Proof. It is well known that we can write

$$A = QDQ^{\mathrm{T}} = \sum_{j=1}^{m} \lambda_j^k v_j v_j^{\mathrm{T}},$$

where Q is an orthogonal matrix with column eigenvectors $\{v_j\}$, and D is a diagonal matrix with entries the eigenvalues $\{\lambda_j\}$. So

$$N_k = \sum_{j=1}^{m} \lambda_j^k u^{\mathrm{T}} v_j v_j^{\mathrm{T}} u$$

and $u^{\mathrm{T}} v_j v_j^{\mathrm{T}} u = (u^{\mathrm{T}} v_j)^2$ with

$$\sum_{j=1}^{m} (u^{\mathrm{T}} v_j)^2 = m.$$

Hence, $N_k = \mathrm{E}(\Lambda^k)$ with $\mathrm{P}(\Lambda = \lambda_j) = (u^{\mathrm{T}} v_j)^2/m$; and, since $\lambda_j > 0\ \forall j$, a consequence of A being positive definite, it is that $\Lambda > 0$ with probability one, and using Jensen's inequality, it is that $\mathrm{E}(\Lambda^k) \geqslant \{\mathrm{E}(\Lambda)\}^k$. So $N_k = m\mathrm{E}(\Lambda^k) \geqslant m\{\mathrm{E}(\Lambda)\}^k = N_1^k/m^{k-1}$, completing the proof. □

Applying Lemma 3.1 shows that

$$\frac{1}{m}\sum_{k=0}^{\infty} f_k N_k t^{-k} \geqslant \sum_{k=0}^{\infty} f_k \frac{N_1^k}{m^k} t^{-k} = f\left(\frac{N_1}{tm}\right).$$

Hence, the inequality (4) is lower bounded by

$$\lambda_{\max}(A) \geqslant tf^{-1}\left(\frac{1}{m}\sum_{k=0}^{\infty} f_k N_k t^{-k}\right) \geqslant tf^{-1}\left(f\left(\frac{N_1}{tm}\right)\right) = \frac{N_1}{m}.$$

In other words, if A is symmetric and positive definite and if $f(x)$ is increasing, then (4) is at least as sharp as the classical bound N_1/m.

A better bound is achieved when all c_k in (5) are made larger than those in (7). This seems possible, because Lemma 3.1 states that the prefactor $N_k/m - (N_1/m)^k$ of f_k/f_1 in (8) is always positive. For, choose $f_2 > 1$, then c_2 is larger. However, increasing f_2 has a negative effect on c_3 since $\frac{N_1^3}{m^3} - \frac{N_1 N_2}{m^2} < 0$. This effect can be compensated by choosing f_3 sufficiently large. A same argument applies for all other terms: there is always the possibility to choose in c_k the highest Taylor coefficient f_k, that is multiplied by $\left(\frac{N_k}{m} - \frac{N^k}{m^k}\right) > 0$, sufficiently large to compensate for the possible decrease in c_k by augmenting lower order Taylor coefficients f_j with $j < k$. It is a matter of optimizing the Taylor coefficients f_k and the bound (6) on t.

4. Numerical examples

Here we consider a specific example when

$$A = \begin{pmatrix} -1 & \sqrt{6} \\ \sqrt{6} & -2 \end{pmatrix}.$$

The eigenvalues of A are 1 and -4 and we have $N_1 = 1.8990$, $N_2 = 2.3031$ and $N_3 = 0.6867$. Hence, the classic bound is given by $N_1/m = 0.9495$. On the other hand, using (7) with $t = 2\widetilde{T}\sqrt{m} = 8.8990$, we obtain a lower bound for $\lambda_{\max}(A)$ as 0.9521, which obviously improves on 0.9495.

Now we consider the example when A is a 10×10 symmetric matrix and for $j = 1, \ldots, i$ we have $a(i, j) = 2j - i$. Then we have $N_1 = 55$, $N_2 = 3553$ and $N_3 = 108823$. Hence the classic bound is given by 5.5. The bound (7) with $t = 2\widetilde{T}\sqrt{10}$, and $\widetilde{T}\sqrt{10} = 50$, is given by the improved lower bound of 9.465. However, for this example, the function $f(x) = x^3$ provides the lower bound of $(N_3/10)^{1/3} = 22.16$.

If we now take $a(i, j) = 2j - 3i$, $j \leqslant i$, and A is again a 10×10 symmetric matrix, then $N_1 = -1375$, $N_2 = 194425$ and $N_3 = -27325375$. Also $\widetilde{T}\sqrt{10} = 190$. So the classic bound is -137.5 and the bound (7) with $t = 2\widetilde{T}\sqrt{10}$ is -136.00. On this occasion the bound based on $f(x) = x^3$ is given by -139.8, which is smaller than the classic bound. The bound (2) is given by -137.00 which improves on the classic bound but is worse than (7).

Appendix A. Taylor expansion of $f^{-1}(\frac{1}{m}\sum_{k=0}^{\infty} f_k N_k z^k)$ around $z = 0$

We now expand $f^{-1}(\frac{1}{m}\sum_{k=0}^{\infty} f_k N_k z^k)$ in a series around $z = \frac{1}{t} = 0$ by invoking characteristic coefficients, defined e.g. in [2, Appendix]. We apply the general expansion (deduced from [2, Appendix]), provided that $f(z_0) = h(z_0)$

$$f^{-1}(h(z)) = z_0 + \sum_{m=1}^{\infty} \frac{h_m(z_0)}{f_1(z_0)}(z-z_0)^m + \sum_{m=2}^{\infty}\sum_{n=2}^{m}\left(\sum_{k=1}^{n-1}(-1)^k \binom{n+k-1}{k} f_1^{-n-k}(z_0)s^*[k,n-1]|_{f(z)}(z_0)\right) \times \frac{s[n,m]|_{h(z)}(z_0)}{n}(z-z_0)^m$$

to $h(z) = \frac{1}{m}\sum_{k=0}^{\infty} f_k N_k z^k = f_0 + \frac{f_1 N_1}{m} z + \frac{1}{m}\sum_{k=2}^{\infty} f_k N_k z^k$ and $z_0 = 0$. Then,

$$f^{-1}(h(z)) = \frac{1}{mf_1}\sum_{k=1}^{\infty} f_k N_k z^k + \sum_{m=2}^{\infty}\sum_{n=2}^{m}\left(\sum_{k=1}^{n-1}(-1)^k \binom{n+k-1}{k} f_1^{-n-k} s^*[k,n-1]|_{f(z)}\right)\frac{s[n,m]|_{h(z)}}{n} z^m.$$

Hence,

$$c_k = \frac{f_k N_k}{mf_1} + \sum_{n=2}^{k}\left(\sum_{j=1}^{n-1}(-1)^j\binom{n+j-1}{j} f_1^{-n-j} s^*[j,n-1]|_{f(z)}\right)\frac{s[n,k]|_{h(z)}}{n}.$$

This can be simplified using $s[k,k] = f_1^k$ and $s[1,m] = f_m$ to explicitly obtain the prefactor of the highest Taylor coefficient f_k in c_k,

$$c_k = \frac{f_k N_k}{mf_1} + \frac{1}{k}\left(\frac{N_1}{m}\right)^k \sum_{j=1}^{k-1}(-1)^j \binom{k+j-1}{j} f_1^{-j} s^*[j,k-1]|_{f(z)} + \sum_{n=2}^{k-1}\left(\sum_{j=1}^{n-1}(-1)^j\binom{n+j-1}{j} f_1^{-n-j} s^*[j,n-1]|_{f(z)}\right)\frac{s[n,k]|_{h(z)}}{n}$$

or

$$c_k = \frac{f_k}{f_1}\left(\frac{N_k}{m} - \left(\frac{N_1}{m}\right)^k\right) + \frac{1}{k}\left(\frac{N_1}{m}\right)^k \sum_{j=2}^{k-1}(-1)^j \binom{k+j-1}{j} f_1^{-j} s^*[j,k-1]|_{f(z)} + \sum_{n=2}^{k-1}\left(\sum_{j=1}^{n-1}(-1)^j\binom{n+j-1}{j} f_1^{-n-j} s^*[j,n-1]|_{f(z)}\right)\frac{s[n,k]|_{h(z)}}{n}. \quad (8)$$

References

[1] X-Q. Liu, T-Z. Huang, Y-D. Fu, A note on improvements on bounds for nonmaximal eigenvalues of symmetric positive matrices, Linear Algebra Appl. 419 (2006) 612–617.

[2] P. Van Mieghem, A new type of lower bound for the largest eigenvalue of a symmetric matrix, Linear Algebra Appl. 427 (2007) 119–129.

[3] S.G. Walker, On recent Cheeger type bounds for non-maximal eigenvalues applied to positive matrices, SIAM J. Matrix Anal. Appl. 25 (2003) 574–581.
[4] S.G. Walker, Improving bounds for nonmaximal eigenvalues of nonnegative matrices, Linear Algebra Appl. 397 (2005) 133–139.