Eigenvalue Estimates for Symmetric Matrices[1]

by

Pablo Tarazaga[2]

Technical Report 87-26, December 1987.

[1]Work supported by CONICET, Republica Argentina.

[2]IMASL, Universidad Nacional de San Luis, San Luis, Argentina and Mathematical Sciences Department, Rice University, Houston, Texas 77251-1892.

1.- INTRODUCTION.

Symmetric and symmetric positive definite matrices have been extensively studied, and there are good characterizations of these sets. We wish to use the setting that in $R^{n\times n}$, the set of symmetric positive semidefinite matrices forms a cone with a very special structure; the identity matrix is the central direction and there exists certain kinds of symmetries around it. The position of each matrix in the cone depends strongly on its eigenvalues and consequently on its rank, we exploit this special structure below.

First, we observe that, when the rank of the symmetric positive semidefinite matrices decrease, then their angles with the identity matrix increase. In this sense rank one matrices are the farthest from the identity, and all of them form a fixed angle with that matrix.

In the final paragraph we obtain the following bounds for the eigenvalues of any symmetric matrix A. If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A and we denote the trace and Frobenius norm by $tr(.)$ and $||.||_F$ respectively, we prove that

$$\left|\frac{tr(A)}{n} - \lambda_i\right| \leq \left[\frac{n-1}{n}\left(||A||_F^2 - \frac{tr(A)^2}{n}\right)\right]^{1/2},$$

for λ_i, $i = 1,\ldots,n$.

Finally one more relation is established for symmetric positive semidefinite matrices, and it concerns the number of eigenvalues above their mean. We show that when the angle between any matrix and the identity increase, then the number of eigenvalues above the mean decreases in the following manner: at most $k-1$ eigenvalues are above the mean if $\dfrac{tr(A)}{||A||_F} < p_k^{1/2}$, where p_k is defined by $p_k^{-1} = \left(\dfrac{n-k+1}{n}\right)^2 + (k-1)\left(\dfrac{1}{n}\right)^2$. Extremal examples of this relation are the identity with all its eigenvalues at the mean and the rank one matrix with only one eigenvalue above the mean. This relation is valid not only for the mean but for any number in the interval $[0, tr(A)]$, then similar results can be obtained.

2.- NOTATION AND FIRST RESULTS.

We denote by S_n the set of symmetric matrices of order n, and by Ω_n the matrices in S_n that are positive semidefinite. Now for k=1,...,n we can define the following subsets

$$\Omega_k^{=} = \{A \in \Omega_n \ / \ rank(A) = k\}$$

and

$$\Omega_k = \{A \in \Omega_n \ / \ rank(A) \leq k\}.$$

We will use in $R^{n\times n}$, the set of square matrices of order n, the Frobenius inner product defined by

$$\langle A, B\rangle_F = tr(A^T B).$$

This inner product allows us to define the cosine of the angle between two matrices in $R^{n \times n}$ by

$$cos(A,B) = \frac{<A,B>_F}{||A||_F \, ||B||_F}.$$

For the 1-norm and 2-norm in R^n we will use the usual notation.

It is well known that Ω_n is a cone and its interior is the set of symmetric positive definite matrices. We show the location of the sets Ω_k , $k \neq n$ in Ω_n with respect to the identity matrix, which will be noted by I. In this sense we have the following simple result.

Lemma 1: If $A \in \Omega_k^=$, then $\frac{tr(A)}{||A||_F} \leq k^{1/2}$.

Proof: Since $A \in \Omega_k^=$, A has k positive eigenvalues $\lambda_1, \ldots, \lambda_k$. We call λ the vector with these components. We need to compute the cosine of the angle formed between A and I as

$$\cos(A,I) = \frac{<A,I>_F}{||A||_F \, ||I||_F} = \frac{tr(A)}{||A||_F \; n^{1/2}}.$$

But using the facts that $tr(A) = ||\lambda||_1$, $||A||_F = ||\lambda||_2$ and the inequality between 1-norm and 2-norm, we have

$$\frac{tr(A)}{||A||_F \; n^{1/2}} = \frac{||\lambda||_1}{||\lambda||_2 \; n^{1/2}} \leq \frac{k^{1/2}}{n^{1/2}},$$

which proves the lemma.

We want to make some observations: first the lemma is valid for $A \in \Omega_k$, second the contrapositive statement gives us a lower bound for the rank of A and finally if $A \in \Omega_k^=$ has all equal eigenvalues, then the equality holds.

Any element in $\Omega_k^=$ can be written as

$$\sum_{i=1}^{k} a_i a_i^T$$

for some $a_i \in R^n$, i=1,...,k.(see [1], [2]). In particular any element in Ω_1 takes the form ss^T for $s \in R^n$. We are interested now in the projection of any element of S_n over the direction generated by elements of Ω_1.

Lemma 2: Let ss^T be an element of Ω_1, the projection of $A \in S_n$ over the direction ss^T is

$$\frac{s^T A s}{s^T s} \frac{ss^T}{s^T s}.$$

Proof: The desired projection is given by

$$||A||_F \; cos(A, ss^T) \; \frac{ss^T}{s^T s}.$$

We can compute now the coefficient of the unitary matrix $\frac{ss^T}{s^T s}$

$$||A||_F \; cos(A,ss^T) = ||A||_F \; \frac{<A,ss^T>_F}{||A||_F \, ||ss^T||_F} = \frac{tr(A^T ss^T)}{||ss^T||_F} = \frac{tr(A^T ss^T)}{s^T s} =$$

$$= \frac{\sum_{i=1}^{n} A_i^T s_i s}{s^T s} = \frac{\sum_{i=1}^{n} s_i (A_i^T s)}{s^T s} = \frac{\sum_{i=1}^{n} s_i (A^T s)_i}{s^T s} = \frac{\sum_{i=1}^{n} s_i (A s)_i}{s^T s} = \frac{s^T A s}{s^T s}.$$

Is is very important to observe that the coefficient computed is the Rayleigh quotient.

3.- MAIN RESULTS.

We consider now the following set

$$\gamma(I; \frac{1}{n^{1/2}}) = \{Y \epsilon R^{n \times n} \ / \ |cos(Y,I)| = \frac{1}{n^{1/2}} \}.$$

For $n=2$ is easy to prove that $\Omega_1 = \gamma(I, \frac{1}{2^{1/2}})$, but for $n>2$ we get

$$\Omega_1 \subset \gamma(I, \frac{1}{n^{1/2}}).$$

Using $\gamma(I, \frac{1}{n^{1/2}})$ allows us to get lower and upper bounds for the eigenvalues that are invariant for similar orthogonal transformations. Second, using this invariance allows us to write the trace and its Frobenius norm in terms of eigenvalues, thus, it is easy to see that

$$[\frac{1}{n} (||A||_F^2 - \frac{tr(A)^2}{n})]^{1/2}$$

is the standard deviation of the eigenvalues λ_i i=1,...,n; this can be computed even if the eigenvalues are unknown.

Theorem 3: Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A \epsilon S_n$, then each λ_i $i=1, \ldots, n$ satisfies

$$|\lambda_i - \frac{tr(A)}{n}| \leq [\frac{n-1}{n} (||A||_F^2 - \frac{tr(A)^2}{n})]^{1/2}.$$

Proof: If λ_n is the maximum eigenvalue of A, then by the properties of the Rayleigh quotient we have

$$\lambda_k = \max_{s \epsilon R^n} \frac{s^T A s}{s^T s} = \max_{s \epsilon R^n} ||A||_F cos(A, ss^T) = \max_{ss^T \epsilon \Omega_1} ||A||_F cos(A, ss^T) \leq \max_{Y \epsilon \gamma(I, \frac{1}{n^{1/2}})} ||A||_F cos(A,Y)$$

where the last inequality is consequence of the inclusion $\Omega_1 \subset \gamma(I, \frac{1}{n^{1/2}})$.

Our goal now is to compute

$$\max_{Y \epsilon \gamma(I, \frac{1}{n^{1/2}})} ||A||_F cos(A,Y).$$

In order to do that, we transform our function as follows

$$||A||_F cos(A,Y) = ||A||_F \frac{<A,Y>}{||A||_F ||Y||_F} = \frac{<A,Y>}{||Y||_F}.$$

We need now a parametrization of the vectors of $\gamma(I,\frac{1}{n^{1/2}})$ in an appropriate way. We propose the following expression for $Y \epsilon \gamma(I,\frac{1}{n^{1/2}})$:

$$Y = \frac{1}{n^{1/2}}\,\frac{I}{n^{1/2}} + \sigma_1\,\frac{\bar{A}}{||\bar{A}||_F} + \sigma_2\,\frac{B}{||B||_F}.$$

In this expression $\bar{A}$ is the projection of A in the subspace orthogonal to I, and B is orthogonal to I and A. In order to get Y unitary, we ask that $\sigma_1^2 + \sigma_2^2 = \frac{n-1}{n}$.

Now we are ready to compute

$$\frac{<A,Y>}{||Y||_F} = \frac{1}{n^{1/2}}\,\frac{<A,I>}{n^{1/2}} + \sigma_1 <A,\frac{\bar{A}}{||\bar{A}||_F}> + \sigma_2 <A,\frac{B}{||B||_F}> =$$

$$= \frac{1}{n} <A,I> + \sigma_1\ ||A||_F \cos(A,\bar{A}) + \sigma_2\, 0 = \frac{1}{n} tr(A) + \sigma_1\,(||A||subAF^2 - \frac{tr(A)^2}{n})^{1/2} + \sigma_2\, 0$$

and is easy to see that

$$\max_{Y\epsilon\gamma(I,\frac{1}{n^{1/2}})} ||A||_F \cos(A,Y) = \max_{\substack{Y\epsilon\gamma(I,\frac{1}{n^{1/2}}) \\ ||Y||_F=1}} <A,Y> = \frac{1}{n}\, tr(A) + (\frac{n-1}{n})^{1/2}\,(||A||_F^2 - \frac{tr(A)^2}{n})^{1/2},$$

because $(||A||_F^2 - \frac{tr(A)^2}{n})^{1/2} \geq 0$.

The lower bound can be computed in a very similar way.

Corollary 4: For $n=2$, these bounds are exactly the eigenvalues.

Corollary 5: In the inequality of the above theorem, the equality holds for $A \epsilon \Omega_1$ or when A is a multiple of the identity matrix.

Another interesting thing to know is if there exist eigenvalues in the intervals

$$[\frac{tr(A)}{n} - (n-1)^{1/2}s\ ,\ \frac{tr(A)}{n} - s] \quad \text{and} \quad [\frac{tr(A)}{n} + s\ ,\ \frac{tr(A)}{n} + (n-1)^{1/2}s].$$

The following result concerns this question.

Corollary 6: At least one of the maximum or the minimum eigenvalue is in one of the intervals mentioned just above.

Proof: It is a consequence of the fact that the standard deviation is less than or equal to the absolute value of the maximum deviation.

Another question that we attack for matrices in Ω_n is, how are the eigenvalues located with respect to their mean? There exists a relation between the angle that the matrix A forms with the identity matrix and the number of eigenvalues above of the mean. The following result establishes this relation. We need to introduce some special values p_k for $k=1,\ldots,n$, defined by

$$p_k^{-1} = (\frac{n-k+1}{n})^2 + (k-1)\,(\frac{1}{n})^2.$$

Theorem 7: If $A \in \Omega_n$ has only k eigenvalues greater or equal to $\frac{tr(A)}{n}$ then

$$\frac{tr(A)}{\|A\|_F} \geq (p_k)^{1/2}.$$

Proof: We can assume that $tr(A)=1$, because this does not affect either the angle between A and I or the order relation between eigenvalues and its mean. We remember that

$$cos(A,I) = \frac{tr(A)}{\|\lambda\|_2 \, n^{1/2}},$$

then it is only necessary to prove that

$$\frac{1}{\|\lambda\|_2} \geq (p_k)^{1/2} \, ,$$

or

$$p_k^{\,-1} \geq \sum_{i=1}^{n} \lambda_i^2 .$$

But this is clear because $(p_k)^{-1}$ is the optimum to the problem

$$\begin{aligned} \max\ & g(\lambda_1, \ldots, \lambda_n) = \sum_{i=1}^{n} \lambda_i^2 \\ \text{s.t.}\ & \sum_{i=1}^{n} \lambda_i = 1 \\ & \lambda_{n-i} \geq \frac{1}{n} \qquad i=1,\ldots,k \\ & \lambda_i \geq 0 \qquad i=1,\ldots,n . \end{aligned}$$

An easy proof of the last statment can be done using the Kuhn-Tucker conditions.

But really, the useful version of this result is the following.

Corollary 8: Let $A \in \Omega_n$. If $\frac{tr(A)}{\|A\|_F} < (p_k)^{1/2}$, then at most $k-1$ eigenvalues are above $\frac{tr(A)}{n}$.

Corollary 9: If $A \in \Omega_n$ and $\frac{tr(A)}{\|A\|_F} < (p_2)^{1/2}$, then A has only one eigenvalue above the $\frac{tr(A)}{n}$ and it belongs to the interval $[\frac{tr(A)}{n} + s(\lambda) \, , \, \frac{tr(A)}{n} + (n-1)^{1/2} s(\lambda)]$.

Finally we want to note, that similar results to theorem 7 and corollary 8 can be obtained with identical proofs for every α in the interval $[0 \, , \, tr(A)]$. The values p_k are now defined by

$$p_k(\alpha)^{-1} = \left(\frac{\alpha-k+1}{\alpha}\right)^2 + (k-1)\left(\frac{1}{\alpha}\right)^2 .$$

All results formulated for symmetric positive semidefinite matrices have the corresponding ones for symmetric negative semidefinite matrices.

ACKNOWLEDGMENTS

I want to express my grateful to Professors John E. Dennis, Beresford N. Parlett and Franchesco Zirilli for their comments and suggestions.

References:

[1] Hill R.D. and S.R. Waters On the Cone of positive Semidefinite Matrices. Linear Algebra and its Applications 90:81-88 (1987)

[2] Tarazaga P. A quadratic minimization problem on subsets of symmetric positive semidefinite matrices.(to appear) 1987.