Structural Bounds for
Eigenvalue Perturbation

Pablo Tarazaga

May, 1991

TR91-13

STRUCTURAL BOUNDS FOR EIGENVALUE PERTURBATION*

by

Pablo Tarazaga†‡

* This work was partially supported by NSF Grant RII-8905080

† Department of Mathematics, University of Puerto Rico, Mayaguez, PR 00709-5000

‡ Part of this work was done while the author was Visiting Professor at Department of Mathematical Sciences, Rice University, Houton, TX 77251-1892

Abastract:

A matrix perturbation $B - A$ in the space of symmetric matrices should be related to the structure of that space. We try to take advantage of this fact to decompose the matrix perturbation in such a way that we get a more precise description of the eigenvalue perturbation. We obtain a lower bound for the eigenvalue perturbation that improves the known bound $| \, \|B\|_F - \|A\|_F \, |$ given by the norm. Also we construct an upper bound that is related to the structure and sometimes is smaller than the known estimate $\|B - A\|_F$. This bound gives us the maximal eigenvalue perturbation of two matrices with the same eigenvectors, keeping the same eigenvalue order.

Keywords: eigenvalue perturbation, matrix perturbation.

AMS(MOS) subject clasification: 15A18, 15A24, 65F15.

1.- INTRODUCTION.

Some properties of the structure of the symmetric matrices and the result obtained in [2] and [3] led us to look at the problem of eigenvalue perturbation from the view point of the relation between the perturbation and the structure of the space. The idea behind this paper is that there are subspaces that preserve the eigenvectors when you move inside them and there are special sets (generated by orthogonal similarity transformations) that preserve eigenvalues in the same way.

These properties suggest one may decompose any matrix perturbation into a pure eigenvalue perturbation plus a pure eigenvector perturbation, as we shall do in section 2.

If B is a perturbation of A we obtain in section 3 a lower bound for the eigenvalue perturbation. We construct a matrix X_* such that

$$\|B - X_*\|_F \leq \Big[\sum_{i=1}^{n}(\mu_i - \lambda_i)^2\Big]^{1/2}$$

where μ_i and λ_i, $i = 1, \ldots, n$ are the eigenvalues of A and B respectively. Moreover this bound satisfies

$$\big|\, \|B\|_F - \|A\|_F \,\big| \leq \|B - X_*\|_F$$

and the equality holds only if B is a scalar multiple of an orthogonal similarity transformation of A. There is a simple construction for X_* using A, B and the identity matrix.

In section 4, under natural order assumptions, we compute an upper bound for the eigenvalue perturbation, obtained via a matrix M which has the eigenvalues of A and the eigenvectors of B. The bound is computed by an optimization approach which yields an estimate for the maximum distance between M and B in the worse case. Sometimes this estimate is worse than the usual $\|B - A\|_F$ estimate (see [1]) and in others it is better. The following example shows the motivation for this estimate.

Define A and B as follows

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}.$$

The eigenvalues of A and B are equal and the vector of eigenvalues is $(1, 0, 0)$ with no eigenvalue perturbation at all, but $\|B - A\|_F = 2^{1/2}(\frac{2}{3})^{1/2}$ while our bound is $(\frac{2}{3})^{1/2}$. This estimate gives us an idea of the maximal perturbation of two matrices with the same eigenvectors, keeping the same eigenvalues order.

Finally in section 4 we obtain a global estimate for the perturbation with no order consideration, and in the same way as we did in section 3, we construct a matrix X^* such that

$$\sum_{i=1}^{n}(\mu_i - \lambda_{\pi(i)})^2 \leq \|B - X^*\|_F^2$$

for any arbitrary permutation π.

2.- PRELIMINARIES.

We denote by $R^{n\times n}$ the space of square matrices of order n and by S_n the subspace of symmetric matrices. We will use in $R^{n\times n}$ the Frobenius inner product defined by

$$<A, B>_F = tr(A^T B)$$

and the Frobenius norm generated by this inner product will be denoted by $\|.\|_F$.

This allows us to introduce the cosine between two matrices as follows

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

If we are interested in the cosine between any matrix and the identity, denoted by I, we have

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

¿From this expression it follows that the cosine between a matrix and the identity is invariant under orthogonal similarity transformations.

Let B be a perturbation of A, both in S_n, and denote the spectral decomposition for these two matrices by

$$A = \sum_{i=1}^{n} \lambda_i v_i v_i^T$$

and

$$B = \sum_{i=1}^{n} \mu_i u_i u_i^T.$$

If P is an orthogonal matrix such that

$$Pu_i = v_i \quad i = 1, \ldots, n,$$

then we are interested in the matrix $P^T AP$ which we denote by M. It is easy to see that M has the same eigenvalues as A and the same eigenvectors as B. This allows us to decompose the perturbation $B - A$ in the following way

$$B - A = (B - M) + (M - A)$$

where $B - M$ is a pure eigenvalue perturbation because both matrices have the same eigenvectors. Also we have that $M - A$ is a pure eigenvector perturbation because both matrices have the same eigenvalues.

A very good question is how to determine whether a perturbation only involves the the eigenvalues or only the eigenvectors. The answer is not easy, but some conditions can be stated.

Lemma 1: A necessary condition for $B - A$ to be an eigenvector perturbation is that

$$tr(A)\|B\|_F = tr(B)\|A\|_F$$

This condition is necessary and sufficient for n=2.

Lemma 2: If $tr(A)\|B\|_F \neq tr(B)\|A\|_F\|$, then the perturbation implies an eigenvalue perturbation. In other words $(B - M) \neq 0$.

The proof of these lemmas are a direct consequence of the invariance of the trace and the Frobenius norm under orthogonal similarity transformations.

We need another piece of notation, given $A \in S_n$ we define

$$\gamma(I, A) = \{Y \in S_n / cos(Y, I) = cos(A, I)\}$$

the conic shell that contains A. As a consequence of our previous discussion, for any orthogonal matrix P we have that

$$P^T AP \in \gamma(I, A).$$

3.- THE LOWER BOUND.

Because the cosine between a matrix A and the identity depends only on the eigenvalues of A, we have that if a matrix X has the same eigenvalues as A then $X \in \gamma(I, A))$. It is well known that $\|B - A\|_F$ is an estimate of the eigenvalue perturbation [1]. But from our decomposition the exact eigenvalue perturbation is given by

$$\|B - M\|_F = \Big[\sum_{i=1}^{n} (\mu_i - \lambda_i)^2\Big]^{1/2},$$

because B and M have the same eigenvectors and the spectral decAomposition is an orthogonal decomposition when we use the Frobenius norm. Then we are looking for a lower bound for $\|B - M\|_F$ that does not require the computation of M, since this would imply that we know the spectral decomposition of A and B.

Because of the facts pointed out above, we want to consider the following optimization problem

$$PLB = \begin{cases} \min \|B - X\|_F^2 \\ X \in \gamma(I, A) \\ \|X\|_F = \|A\|_F \end{cases}$$

which will provide us with a lower bound for our eigenvalue perturbation. It is clear that this is a lower bound for $\|B - A\|_F$ and $\|B - M\|_F$.

In order to solve PLB we will search for a matrix X that

a) belongs to $\gamma(I, A)$

b) lies in the plane generated by I and B

c) has Frobenius norm equal to $\|A\|_F$.

We will do this by first generating a matrix in the plane of I and B and then force it to satisfies a) and c).

Because of symmetric consideration our matrix must be $B + \nu I$ where ν is a real number. We are going to determine ν from the condition a) or in other words our matrix must satisfy

$$cos(I, B + \nu I) = cos(I, A)$$

or

$$(1) \qquad \frac{tr(B) + \nu n}{\|B + \nu I\|_F} = \frac{tr(A)}{\|A\|_F}$$

If we square the equality we get the following quadratic equation for ν

$$[n(n\|A\|_F^2 - tr(A)^2)]\nu^2 + [2tr(B)(n\|A\|_F^2 - tr(A)^2)]\nu + [tr(B)^2\|A\|_F^2 - \|B\|_F^2 tr(A)^2] = 0$$

and after some algebraic computations we obtain

$$\nu = -\frac{tr(B)}{n} \pm \frac{tr(A)}{n}\left(\frac{n\|B\|_F^2 - tr(B)^2}{n\|A\|_F^2 - tr(A)^2}\right)^{1/2}.$$

If $X \in S_n$, then the standard deviation of the eigenvalues of X can be written as

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

which will be denoted by $s(X)$. Then

$$ns(X) = \left[(n\|X\|_F^2 - tr(X)^2)\right]^{1/2}$$

which allows us to change the expression of ν as follows

$$\nu = -\frac{tr(B)}{n} \pm \frac{tr(A)}{n}\frac{s(B)}{s(A)}.$$

Since $s(B)$ and $s(A)$ are always nonnegative and since ν has to satisfy the equation (1), it is easy to see that the negative sign in the second term of ν does not provide a solution (this is an extraneous solution that appeared because we squared equation (1)). Hence

$$(2) \qquad \nu = -\frac{tr(B)}{n} + \frac{tr(A)}{n}\frac{s(B)}{s(A)}.$$

Remark: In [2] we proved that for any symmetric matrix Z, all its eigenvalues are in the interval

$$\left[\frac{tr(Z)}{n} - [\frac{n-1}{n}(\|Z\|_F^2 - \frac{tr(Z)^2}{n})]^{1/2}, \ \frac{tr(Z)}{n} + [\frac{n-1}{n}(\|Z\|_F^2 - \frac{tr(Z)^2}{n})]^{1/2}\right].$$

Thus the length of the interval that contains the eigenvalues is

$$2\left[\frac{n-1}{n}\left(\|Z\|_F^2 - \frac{tr(Z)^2}{n}\right)\right]^{1/2}.$$

Now it is interesting to observe that

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

which means that $s(B)/s(A)$ is the quotient between the lengths of the intervals that contain the eigenvalues of A and B.

Finally we obtain the solution of our optimization problem, which is

$$X_* = \frac{\|A\|_F}{\|B+\nu I\|_F}(B+\nu I)$$

where ν is given in (2). Using equation (1) and the expression of ν we have

$$\frac{\|A\|_F}{\|B+\nu I\|_F} = \frac{s(A)}{s(B)}$$

$$X_* = \frac{s(A)}{s(B)}(B+\nu I) = \frac{s(A)}{s(B)}B + \left[\frac{tr(A)}{n} - \frac{s(A)}{s(B)}\frac{tr(B)}{n}\right]I.$$

These lemmas allow us to establish the following result.

Lemma 3: X_* solves the optimization problem PLB, and

$$\min_{\substack{X\in\gamma(I,A)\\ \|X\|_F=\|A\|_F}} \|X-B\|_F = \|X_*-B\|_F.$$

Proof: The conditions a), b) and c) are the Kuhn-Tucker conditions.

With this result we can state the main result of this section.

Theorem 4: Given a perturbation $B \in S_n$ of $A \in S_n$ then

$$\|B - X_*\|_F \le \|B - A\|_F$$

and

$$\|B - X_*\|_F \le \|B - M\|_F = \left[\sum_{i=1}^{n}(\mu_i - \lambda_i)^2\right]^{1/2}.$$

We want to know the relation between the lower bound $\|B-X_*\|_F$ and the well known bound $|\ \|B\|_F - \|A\|_F|$. Because $\|A\|_F = \|X_*\|_F$, then

$$|\ \|B\|_F - \|A\|_F| = |\ \|B\|_F - \|X_*\|_F| \le \|B - X_*\|_F.$$

Corollary 5: Given a perturbation B of A, then

$$|\ \|B\|_F - \|A\|_F| \le \|B - X_*\|_F$$

and equality holds if and only if B is a multiple of P^TAP for some orthogonal matrix P.

4.- UPPER BOUNDS.

We can assume that the eigenvalues of A and its perturbation B are ordered as follow

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$$

$$\mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$$

which implies M has the same order as A for its eigenvalues.

Our principal problem now is to determine the maximum distance between M and B. It is important to note that they are in the same subspace (generated by the rank one matrices $u_i u_i^T$ $i = 1, \ldots, n$).

Our problem can be stated as a constrained optimization problem, which we will call $P1$

$$\max \|X - Y\|_F^2 = \sum_{i=1}^{n} (\eta_i - \sigma_i)^2$$

$$\Gamma(M) \begin{cases} \sum_{i=1}^{n} \sigma_i = tr(M) \\ \sum_{i=1}^{n} \sigma_i^2 = \|M\|_F^2 \\ \sigma_i - \sigma_{i-1} \geq 0 \quad i = 1, \ldots, n \end{cases}$$

$$\Gamma(B) \begin{cases} \sum_{i=1}^{n} \eta_i = tr(B) \\ \sum_{i=1}^{n} \eta_i^2 = \|B\|_F^2 \\ \eta_i - \eta_{i-1} \geq 0 \quad i = 1, \ldots, n \end{cases}$$

where we are assuming for X and Y the following spectral representation

$$X = \sum_{i=1}^{n} \sigma_i u_i u_i^T$$

$$X = \sum_{i=1}^{n} \eta_i u_i u_i^T.$$

Notice first that $\Gamma(M) = \Gamma(A)$ and second that this problem is a generalization of the LB problem used in [3].

This problem has a very special structure. In general $\Gamma(M)$ and $\Gamma(B)$ are different sets, but when $tr(A) = tr(B)$ and $\|A\|_F = \|B\|_F$ they become the same set (this particular case is important in the proof). Each of these sets $\Gamma(.)$ are "spheric polytopes", in the sense that all the constraints are linear except the equation related to the norm, which is quadratic.

Our first goal is to identify the extremal points of these sets. Because they are structurally equivalent, we only establish the result for one of them.

Lemma 6: The extremal points of $\Gamma(M)$ are the matrices $X(k)$ $k = 1, \ldots, n-1$, whose eigenvalues are

$$\sigma_i = \frac{tr(M)}{n} - \left[\frac{n-k}{nk}\left(\|A\|_F^2 - \frac{tr(M)^2}{n}\right)\right]^{1/2} \quad i = 1, \ldots, k$$

$$\sigma_i = \frac{tr(M)}{n} + \left[\frac{k}{n(n-k)}\left(\|A\|_F^2 - \frac{tr(M)^2}{n}\right)\right]^{1/2} \quad i = k+1, \ldots, n.$$

Note that because M and B are in the same spectral subspace, then it is possible to construct $X(k),\ k = 1, \ldots, n-1$ from its eigenvalues using the eigenvectors of B.

Proof: Because we need n equations from $\Gamma(M)$ in order to find the extremal points and we have two equations and $n-1$ inequalities, we transform $n-2$ of the inequalities into equations. Because these new equations come from the order inequalities, we have to set the first k eigenvalues equal to α and the remaining $n-k$ equal to β. With this selection for $k = 1, \ldots, n-1$ the inequality $\sigma_{k+1} - \sigma_k \geq 0$ remains as the only possible strict inequality. These considerations give us a new system to solve

$$\begin{cases} k\alpha + (n-k)\beta = tr(M) \\ k\alpha^2 + (n-k)\beta^2 = \|M\|_F^2. \end{cases}$$

After some elementary algebra we get

$$\alpha = \frac{tr(M)}{n} \pm \left[\frac{n-k}{nk}\left(\|A\|_F^2 - \frac{tr(M)^2}{n}\right)\right]^{1/2}$$
$$\beta = \frac{tr(M)}{n} \pm \left[\frac{k}{n(n-k)}\left(\|A\|_F^2 - \frac{tr(M)^2}{n}\right)\right]^{1/2}.$$

It is easy to see that the second terms of α and β must have opposite signs in order to satisfy the equation of $\Gamma(M)$ that includes $tr(M)$. Finally the increasing order of the eigenvalues tells us that the sign of α must be negative and positive for β. Hence for $k = 1, \ldots, n-1$ the matrix $X(k)$ has eigenvalues

$$\sigma_i = \begin{cases} \alpha & i = 1, \ldots, k \\ \beta & i = k+1, \ldots, n. \end{cases}$$

with

$$\alpha = \frac{tr(M)}{n} - \left[\frac{n-k}{nk}\left(\|A\|_F^2 - \frac{tr(M)^2}{n}\right)\right]^{1/2}$$
$$\beta = \frac{tr(M)}{n} + \left[\frac{k}{n(n-k)}\left(\|A\|_F^2 - \frac{tr(M)^2}{n}\right)\right]^{1/2}.$$

In order to simplify these expressions, we introduce the notation

$$\delta(k) = \left[\frac{n-k}{k}\right]^{1/2},$$

and observe Bœ[B

$$\delta(n-k) = \frac{1}{\delta(k)}.$$

This allows us to express σ_i as

$$\sigma_i = \frac{tr(M)}{n} - \delta(k)s(M) \quad i = 1, \ldots, k$$
$$\sigma_i = \frac{tr(M)}{n} + \delta(n-k)s(M) \quad i = k+1, \ldots, n$$

We denote by $\lambda[X(k)]$ the vector of eigenvalues of $X(k)$. In other words the vector with components σ_i.

In the same way the extreme points of $\Gamma(B)$ are matrices $Y(j),\ j = 1, \dots, n-1$ whose eigenvalues are

$$\eta_i = \begin{cases} \theta & i = 1, \dots, j \\ \epsilon & i = j+1, \dots, n. \end{cases}$$

where

$$\theta = \frac{tr(B)}{n} - \delta(j)s(B)$$
$$\epsilon = \frac{tr(B)}{n} + \delta(n-j)s(B).$$

Taking into account that B and M are in the same subspace we have

$$\|X(k) - Y(j)\|_F = \|\lambda[X(k)] - \lambda[Y(j)]\|_2$$

and this distance will be denoted by $d(k, j)$.

An equivalent optimization problem to $P1$ is the following problem $P2$

$$\max_{\substack{k=1,\dots,n-1 \\ j=1,\dots,n-1}} d(k,j) = \|\lambda[X(k)] - \lambda[Y(j)]\|_2.$$

Before considering this new problem, we write down the vector $\lambda[X(k)] - \lambda[Y(j)]$. Assuming that $k \leq j$,

$$(\lambda[X(k)] - \lambda[Y(j)])_i = \begin{cases} \frac{tr(M)}{n} - \frac{tr(B)}{n} - \delta(k)s(M) + \delta(j)s(B) & 1 \leq i \leq k \\ \frac{tr(M)}{n} - \frac{tr(B)}{n} + \delta(n-k)s(M) + \delta(j)s(B) & k+1 \leq i \leq j \\ \frac{tr(M)}{n} - \frac{tr(B)}{n} + \delta(n-k)s(M) - \delta(n-j)s(B) & j+1 \leq i \leq n. \end{cases}$$

First we are going to consider a special but basic case of $P2$, where $\|M\|_F = \|B\|_F$ and $tr(M) = tr(B)$. We wish to compute $d(k, j)$ for this case.

$$d(k,j)^2 = \Big\{ k[\delta(j) - \delta(k)]^2 + (j-k)[\delta(n-k) + \delta(j)]^2$$
$$+ (n-j)[\delta(n-k) - \delta(n-j)]^2 \Big\} s(B)^2.$$

If we expand the squared terms and regroup we obtain

$$d(k,j)^2 = 2s(B)^2 \left[n - \frac{kn}{n-k} \delta(j)\delta(k) \right].$$

This problem has certain symmetries among the distances that we can exploit.

Lemma 7: $d(k,j) = d(n-k, n-j)$.

Proof: We only need to compute $d(n-k,n-j)$. Assuming for simplicity that $k \le j$ and using the symmetry of the distance we obtain

$$d(n-k,n-j)^2 = d(n-j,n-k)^2 = \Big\{(n-j)[\delta(n-k)-\delta(n-j)]^2 + [(n-k)-(n-j)][\frac{1}{\delta(n-j)}+\delta(n-k)]^2 + [n-(n-k)][\frac{1}{\delta(n-j)}-\frac{1}{\delta(n-k)}]^2\Big\}s(B)^2.$$

But using the definition of $\delta(.)$ and regrouping, we obtain

$$\Big\{k[\delta(j)-\delta(k)]^2+(j-k)[\delta(n-k)+\delta(j)]^2+(n-j)[\delta(n-k)-\delta(n-j)]^2\Big\}s(B)^2 = d(k,j)^2.$$

We want to make a couple of observations. First we only need to consider $k \le j$ because of the symmetry of the distance. Second, it is enough to consider $k+j \le n$ because if $k+j > n$ then $(n-k)+(n-j) \le n$ and these distances are equal by Lemma 6. Moreover the case $k=j$ gives us distances equal to zero, and this is not important for our optimization problem, which is a maximization problem. These observations introduce new important constraint in our problem.

There is an interesting subset of distances which we call maximal distances. They are the distances $d(k,j)$ with the property that $k+j=n$, or in other words, distances with the form $d(k,n-k)$. A useful property of this set of distances is our next result.

Lemma 8: If $\bar{k} \le k$, then

$$d(k,n-k) \le d(\bar{k},n-\bar{k})$$

and $d(1,n-1)$ is the maximum of the set.

Proof: Using the formula for the distance, we can compute

$$d(k,n-k)^2 = 2s(B)^2[n-\frac{kn}{n-k}\delta(n-k)\delta(k)] = 2s(B)^2(n-\frac{kn}{n-k}).$$

Now from elemental calculus, it is easy to verify that this function decreases with k.

Our next step is to prove that given any distance $d(k,j)$, it is always bounded by a maximal distance.

Lemma 9: Given $d(k,j)$, if $i=\min\{k,n-k\}$, then

$$d(k,j) \le d(i,n-i).$$

Proof: Because of the symmetry of the distance, we can always suppose that i is equal to k. Then we have to prove

$$d(k,j) \le d(k,n-k)$$

or

$$2s(B)^2[n-\frac{nk}{n-k}\delta(k)\delta(j)] \le 2s(B)^2[n-\frac{nk}{n-k}],$$

and the inequality follows from the fact that $\delta(k)\delta(j) \geq 1$ under the condition $k + j \leq n$.

These lemmas establish the following result.

Theorem 10: If $\|M\|_F = \|B\|_F$ and $tr(M) = tr(B)$, then

$$\max_{\substack{k=1,\ldots,n-1 \\ j=1,\ldots,n-1}} d(k,j) = d(1, n-1).$$

In others words, $d(1, n-1)$ solves $P2$ and $P1$ under the hypothesis given above.

We now wish to solve problem $P1$ in general and we need to modify some notation. We will use ϕ to denote

$$\frac{tr(M) - tr(B)}{n}$$

and e to denote the vector with all components equal to one. Then we can write

$$\lambda[X(k)] - \lambda[Y(j)] = \phi e + z(k,j)$$

where

$$z(k,j) = \begin{cases} \delta(j)s(B) - \delta(k)s(M) & 1 \leq i \leq k \\ \delta(n-k)s(M) + \delta(j)s(B) & k \leq i \leq j \\ \delta(n-k)s(M) - \delta(n-j)s(B) & j \leq i \leq n \end{cases}.$$

Because $\lambda[X(k)]$ and $\lambda[Y(j)]$ are vectors of eigenvalues satisfying $\Gamma(M)$ and $\Gamma(B)$, then

$$\lambda[X(k)]^T e = tr(M)$$

$$\lambda[Y(j)]^T e = tr(B)$$

or

$$\{\lambda[X(k)] - \lambda[Y(j)]\}^T e = tr(A) - tr(B)$$

for $k = 1, \ldots, n-1$ and $j = 1, \ldots, n-1$. But using our new notation

$$[\phi e + z(k,j)]^T e = \phi n + z(k,j)^T e = tr(M) - tr(B)$$

we have that

$$z(k,j)^T e$$

is constant for any k and j. But we know that

$$\|\lambda[X(k)] - \lambda[Y(j)]\|_2^2 = \|\phi e + z(k,j)\|_2^2 = \|\phi e\|_2^2 + \|z(k,j)\|_2^2 + 2\phi e^T z(k,j).$$

Now it is easy to see that our goal is to maximize $d(k,j)^2$ but it is sufficient to maximize $\|z(k,j)\|_2^2$ in order to obtain the maximizer.

We want to look for a useful expression for $\|z(k,j)\|_2^2$, since

$$\|z(k,j)\|_2^2 = k[\delta(j)s(B) - \delta(k)s(M)]^2 + (j-k)[\delta(n-k)s(M) + \delta(j)s(B)]^2 + (n-j)[\delta(n-k)s(M) - \delta(n-j)s(B)]^2.$$

By expanding the squares and regrouping, we obtain

$$\|z(k,j)\|_2^2 = n[s(B)^2 + s(M)^2] - 2\frac{kn}{n-k}\delta(j)\delta(k)s(B)s(M).$$

Since this is very similar to the function $d(k,j)$, it is easy to define the maximal distances $\|z(k,n-k)\|_2$ and prove results similar to Lemma 7, 8 and 9, which allow us to state the next result.

Theorem 11: The vectors $\lambda[X(1)]$ and $\lambda[Y(n-1)]$ solve problem $P1$ and the maximum value of the objective function is $\|\lambda[X(1)] - \lambda[Y(n-1)]\|_2$.

This completes our results about problem $P1$, and now we can establish the main result of this section.

Theorem 12: Given $B \in S_n$ a perturbation of $A \in S_n$, then

$$\sum_{i=1}^{n}(\lambda_i - \mu_i)^2 \leq (\alpha-\theta)^2 + (n-1)(\beta-\theta)^2 + (\beta-\epsilon)^2,$$

where

$$\alpha = \frac{tr(A)}{n} - \delta(1)s(A)$$
$$\beta = \frac{tr(A)}{n} + \delta(n-1)s(A)$$

and

$$\theta = \frac{tr(B)}{n} - \delta(n-1)s(B)$$
$$\epsilon = \frac{tr(B)}{n} + \delta(1)s(B)$$

Finally we consider a global problem

$$\max_{\substack{X\in\gamma(I,A)\\ \|X\|_F=\|A\|_F}} \|B - X\|_F.$$

Taking into account that the constraints can be described as

$$tr(X) = tr(A)$$

$$\|X\|_F = \|A\|_F,$$

this problem is very similar to the problem discussed in section 3, and the same techniques can be applied to find the solution. These considerations yield to the following result.

Theorem 13: If $B \in S_n$ is a perturbation of $A \in S_n$ and π a permutation of n elements, then

$$\sum_{i=1}^{n}(\mu_i - \lambda_{\pi(i)})^2 \leq \|B - X^*\|$$

where

$$X^* = \frac{s(A)}{s(B)}(-B + \xi I)$$

with

$$\xi = \frac{tr(B)}{n} + \frac{tr(A)}{n}\frac{s(B)}{s(A)}.$$

Proof: Consider that any matrix with the same eigenvalues of A is in $\gamma(I, A)$, and take the farthest one from B, using the plane that contains B and the identity matrix.

5.- CONCLUSIONS.

Even though the results of this paper have been stated for symmetric matrices, they can be generalized in a strightforward manner to Hermitian matrices and the proofs are exactly the same.

Only a few results can be generalized beyond the Hermitian matrices case. The lower bound and the global upper bound still hold for normal matrices, because the tools that we have used to prove these result are the invariance of the trace and the Frobenius norm under similar orthogonal transformations.

Finally we want to say that intermediate problems between Theorem 10 and the global upper bound can be studied, allowing certain shifts in the eigenvalues order. The way to do this is to establish the corresponding optimization problem and solve it in order to obtain the bounds.

REFERENCES:

[1] Horn R.A. and C.R. Johnson, Matrix Analisys, Cambridge University Press 1985.

[2] Tarazaga P., Eigenvalue Estimates for Symmetric Matrices, Linear Algebra and its Applications 135:171-179 (1990).

[3] Tarazaga P., More Estimates for Eigenvalues and Singular Values, Linear Algebra and its Applications 149:97-110 (1991).