UNCLASSIFIED

278 072

Reproduced
by the

ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA
NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.
Enclosure Theorems for Eigenvalues and Probability Inequalities

April 1962

Albert W. Marshall
Ingram Olkin
Mathematics Research
ENCLOSURE THEOREMS FOR EIGENVALUES
AND PROBABILITY INEQUALITIES*

by

Albert W. Marshall

and

Ingram Olkin
Stanford University

Mathematical Note No. 255
Mathematics Research Laboratory
BOEING SCIENTIFIC RESEARCH LABORATORIES
April 1962

*This report is also being issued as Technical Report No. 1, prepared under the auspices of National Science Foundation Grant 19138, Applied Mathematics and Statistics Laboratories, Stanford University, Stanford, California, May 7, 1962.
ENCLOSURE THEOREMS FOR EIGENVALUES
AND PROBABILITY INEQUALITIES

by
Albert W. Marshall and Ingram Olkin

Boeing Scientific Research Laboratories
and
Stanford University

1. Introduction.

Let \( H \) be an Hermitian matrix, \( z \) a row vector of unit length, \( \mu = zHz^* \), \( \sigma^2 = z(H - \mu I)^2z^* \). The enclosure theorem of Krylov and Bogoliubov [1] (see also Weinstein [2], Kohn [3], Block and Fuchs [4], Swanson [5]) asserts that there is an eigenvalue in the interval \([\mu - \sigma, \mu + \sigma]\).

We now present a probabilistic interpretation which yields a simple proof of this result as well as other enclosure theorems.

Write \( H = U^*DU \), where \( U \) is unitary, \( D = \text{diag}(\theta_1, \ldots, \theta_k) \), then for any unit vector \( z \),

\[
\mu = zHz^* = (zU^*)D(Uz^*) = \sum p_j \theta_j,
\]

with \( 0 \leq p_j \), \( \sum p_j = 1 \). Consequently, we may view \( \mu \) as the mean of a random variable \( X \) with probability distribution \( P(X = \theta_j) = p_j, j = 1, \ldots, k \). Similarly, \( \sigma^2 = z(H - \mu I)^2z^* = \sum p_j(\theta_j - \mu)^2 \) is the variance of \( X \). According to the Bienaymé-Chebyshev inequality,

\[
\frac{1}{\mathbf{The work of the second author was supported in part by the National Science Foundation.}}
\]
Hence, for all \( t > 1 \) there is positive mass in the interval \([\mu - \sigma t, \mu + \sigma t]\), which implies that there is positive mass in \([\mu - \sigma, \mu + \sigma]\).

Since the mass is distributed only on eigenvalues, this means that there is at least one eigenvalue in \([\mu - \sigma, \mu + \sigma]\).

In general, if a bound is obtained for some set \( \mathcal{S} \), i.e.,

\[
P(\mathbf{X} \in \mathcal{S}) \geq p,
\]

then \( \sum_{j} p_{j} \geq p \), where the summation is over all \( j \) corresponding to \( \theta_{j} \in \mathcal{S} \). But \( p_{j} = |u_{j} z^{*}|^{2} \), where \( u_{j} \) is the \( j \)-th row of \( U \), and is, therefore, an eigenvector corresponding to the eigenvalue \( \theta_{j} \), so that \( \sum_{\mathcal{S}} |u_{j} z^{*}|^{2} \geq p \).

In Section 2 we state a number of enclosure theorems, and in Section 3 we present the related probabilistic inequalities which furnish proofs.

2. **Enclosure Theorems.**

We now state a number of enclosure theorems to illustrate the variety of results which may be obtained in this manner.

In the following we let \( \mathbf{H} \) be an Hermitian matrix, \( z \) a row vector of unit length, \( \mu = z \mathbf{H} z^{*} \), \( \sigma^{2} = z(\mathbf{H} - \mu \mathbf{I}) z^{*} \).

2.1 **General Hermitian Matrix.**

There exists at least one eigenvalue in the interval

\[
\left[ a + \mu - \sqrt{\sigma^{2} + a^{2}}, a + \mu + \sqrt{\sigma^{2} + a^{2}} \right].
\]
Remark.

In this and the following theorems, a slightly stronger version can be obtained by making the interval half-open.

This theorem is given by Kato [6]. The result of Section 1 is a special case with \( a = 0 \).

2.2 Positive Definite Hermitian Matrix.

Let \( H \) be positive definite, and \( \mu_{-1} = z H^{-1} z^* \). If \( s < \mu \), \( s < 1/\mu_{-1} \), then the interval \([s, (\mu - s) / (1 - s \mu_{-1})]\) contains at least one eigenvalue.

The particular interval, \([(1 - \sqrt{\mu_{-1}} - 1) / \mu_{-1}, (1 + \sqrt{\mu_{-1}} - 1) / \mu_{-1}]\), has minimum length.

2.3 Bounded Hermitian Matrix.

Let \( H \) and \( I - H \) be positive definite. If \( -\mu \leq r \leq -\sigma^2 / (1 - \mu) \), then there is at least one eigenvalue in \([0, r + \mu] \cup [-\sigma^2 / r + \mu, 1]\).

3. Probability Inequalities.

We now present the probability inequalities which furnish proofs for the enclosure theorems of Section 2.

3.1 Theorem of Selberg [7].

If \( X \) is a random variable with \( EX = \mu \), \( E(X - \mu)^2 = \sigma^2 \), then for all real \( a \) and positive \( b \),

(i) \( P( |X - \mu - a| < b ) \geq 1 - \frac{\sigma^2 + a^2}{b^2} \), if \( \sigma^2 > |a| b - a^2 \),

(ii) \( P( |X - \mu - a| < b ) \geq 1 - \frac{\sigma^2}{\sigma^2 + (b - |a|)^2} \), if \( \sigma^2 < |a| b - a^2 \).

This inequality is sharp.
The proof of this asymmetric version of the Bienaymé-Chebyshev inequality in [7] is unnecessarily complicated, and we give the following proof.

**Proof.**

If \( f(x) = \frac{(x - a)^2}{b^2} \) whenever \( \sigma^2 \geq |a| b - a^2 \), \( f(x) = \frac{[(b - |a|) x - \sigma^2]^2}{[(b - |a|)^2 + \sigma^2]^2} \) whenever \( \sigma^2 < |a| b - a^2 \), then \( f(x) \geq 0 \) for all \( x \) and \( f(x) \geq 1 \) for all \( |x - \mu - a| \geq b \).

The result follows from \( E f(X) \geq P( |X - \mu - a| \geq b ) \).

If \( b^2 - a^2 \geq \sigma^2 \geq |a| b - a^2 \), the extremal distribution is

\[
P(Y - \mu = a + b) = p_1, \quad P(Y - \mu = a - b) = p_2, \quad P(Y - \mu = a) = 1 - p_1 - p_2,
\]

where \( p_1 = (\sigma^2 - ab + a^2) / 2b^2 \), \( p_2 = (\sigma^2 + ab + a^2) / 2b^2 \).

If \( \sigma^2 < |a| b - a^2 \), \( a > 0 \), the extremal distribution is

\[
P(Y - \mu = a - b) = \sigma^2 [ (b - a)^2 + \sigma^2 ]^{-1} = 1 - P(Y - \mu = \sigma^2 / (b - a)).
\]

For \( a < 0 \), use \(-Y\) for \( Y \).

If \( \sigma^2 > b^2 - a^2 \), the extremal distribution is

\[
P(Y - \mu = a + b - c) = (b + c - a) / [ 2(b + c) ] = 1 - P(Y - \mu = a - b - c), \quad \text{where}
\]

\[c = \sqrt{\sigma^2 + a^2} - b.\]

To obtain the enclosure theorem, we consider (i) and require \((\sigma^2 + a^2) / b^2 < 1\), so that for all \( b > \sqrt{\sigma^2 + a^2} \), there is an eigenvalue in the interval \([ a + \mu - \sqrt{\sigma^2 + a^2}, a + \mu + \sqrt{\sigma^2 + a^2} ]\).

Since the bound (ii) is always less than 1, we obtain the shortest interval by choosing \( b = (\sigma^2 + a^2) / |a| \). But this leads to the interval \([ a + \mu - (\sigma^2 + a^2) / |a|, a + \mu + (\sigma^2 + a^2) / |a| ]\) which is contained in \([ a + \mu - \sqrt{\sigma^2 + a^2}, a + \mu + \sqrt{\sigma^2 + a^2} ]\).

The extremal distribution for (i) is unique and places positive probability at both of the points \( a + \mu - b \) and \( a + \mu + b \). If we
eliminate this extremal distribution, the probability is strictly less than the bound, so that \( P( |X - \mu - a| < \sqrt{\sigma^2 + a^2} ) > 0 \), which says that there is an eigenvalue in the open interval \( (a + \mu - \sqrt{\sigma^2 + a^2}, a + \mu + \sqrt{\sigma^2 + a^2}) \). By using a half-open interval, the extremal distribution is included.

3.2 An Inequality with Negative Moments.

If \( X \) is a positive random variable, i.e., \( P(X \leq 0) = 0 \), \( EX = \mu \), \( EX^{-1} = \mu_{-1} \), then

\[
P( s < X < t ) \geq 1 - \frac{\mu - 2 \sqrt{st} + st\mu_{-1}}{(\sqrt{t} - \sqrt{s})^2}.
\]

Equality is achieved if and only if the bound is non-negative and \( X \) has the distribution \( P(X = s) = (a + bt) \sqrt{s} / (t - s) \), \( P(X = t) = (a + bs) \sqrt{t} / (t - s) \), \( P(X = \sqrt{st}) = 1 - (a + b \sqrt{st}) / (\sqrt{t} - \sqrt{s}) \), where \( a = (\mu - \sqrt{st}) / (\sqrt{t} - \sqrt{s}) \), \( b = (\mu_{-1} \sqrt{st} - 1) / (\sqrt{t} - \sqrt{s}) \).

Proof.

If \( f(x) = (x - \sqrt{st})^2 / [x (\sqrt{t} - \sqrt{s})^2] \), then \( f(x) \geq 0 \) for all \( x > 0 \), \( f(x) \geq 1 \) for all \( x \notin (s, t) \), and the result follows from \( E f(X) \geq P(X \notin (s, t)) \). The extremal distribution is then determined by placing probability where \( f(x) \) is zero or one.

We note that \( \mu_{-1} \geq 1 \), by the Schwarz inequality, so that \( s < \mu < t, s < 1/\mu_{-1} < t \), follow from the condition that the bound be non-negative.

To obtain the corresponding enclosure theorem, note that the bound is non-negative when \( t \geq (\mu - s) / (1 - \mu_{-1}s) \).
3.3 **Bounded Random Variables.**

If \( X \) is a bounded random variable, i.e., \( P(\alpha \leq X - \mu \leq \beta) = 1 \), \( \mathbb{E}(X - \mu)^2 = \sigma^2 \), \( \alpha \leq s \leq t \leq \beta \), then \( P(X < s \text{ or } X > t) \geq p \).

given as follows:

(1) \( 1 + \frac{\alpha \beta + \sigma^2}{(s-\alpha)(\beta-s)} \), if \( s + t \leq \alpha + \beta \), \( \sigma^2 \geq -\alpha s, \sigma^2 \geq -\beta s \),

(ii) \( 1 + \frac{\alpha \beta + \sigma^2}{(t-\alpha)(\beta-t)} \), if \( s + t \geq \alpha + \beta \), \( \sigma^2 \geq -\alpha t, \sigma^2 \geq -\beta t \),

(iii) \( \frac{st + \sigma^2}{(\beta-s)(\beta-t)} \), if \( s + t \leq \alpha + \beta \), \( -\beta t \leq \sigma^2 \leq -\beta s \), \( \sigma^2 \geq -st \),

(iv) \( \frac{st + \sigma^2}{(s-\alpha)(\beta-t)} \), if \( s + t \geq \alpha + \beta \), \( -\alpha s \leq \sigma^2 \leq -\alpha t \), \( \sigma^2 \geq -st \),

(v) \( \frac{\sigma^2}{s^2 + \sigma^2} \), if \( \sigma^2 < -\alpha s \),

(vi) \( \frac{t^2}{t^2 + \sigma^2} \), if \( \sigma^2 < -\beta t \),

(vii) \( 0 \), otherwise.

This inequality is sharp.

**Proof.**

Write \( y = x - \mu \) and let

(1) \( f(y) = (y - \alpha)(y - \beta) / [(s - \alpha)(\beta - s)] \),

(ii) \( f(y) = -(y - \alpha)(y - \beta) / [(\beta - t)(t - \alpha)] \),

(iii) \( f(y) = 1 - (y - s)(y - t) / [(\beta - s)(\beta - t)] \),

(iv) \( f(y) = 1 - (y - s)(y - t) / [(t - \alpha)(s - \alpha)] \),

(v) \( f(y) = (\alpha y + \sigma^2)^2 / (s^2 + \sigma^2)^2 \),

(vi) \( f(y) = (\beta y + \sigma^2)^2 / (t^2 + \sigma^2)^2 \),

6
then, for each respective case, \( f(y) \geq 0 \) for all \( x \in [\alpha, \beta] \), and \( f(y) \geq 1 \) for all \( x \in [s, t] \), so that \( E f(X) \geq P(X \notin [s, t]) \).

Sharpness is exhibited by the following examples. Let \( Z \) be a random variable with the distribution

\[
P(Z - \mu = a_1) = \frac{(\sigma^2 + a_2a_3)}{[(a_2 - a_1)(a_2 - a_1)]},
\]

\[
P(Z - \mu = a_2) = \frac{(\sigma^2 + a_1a_3)}{[(a_2 - a_3)(a_2 - a_1)]},
\]

\[
P(Z - \mu = a_3) = \frac{(\sigma^2 + a_1a_2)}{[(a_3 - a_2)(a_3 - a_1)]}.
\]

The hypotheses are satisfied and equality is attained in cases (i) - (iv) if \( (a_1, a_2, a_3) \) is taken to be: \( (\alpha, t, \beta) \) for (i), \( (s, t, \beta) \) for (ii), \( (\beta, t, \alpha) \) for (iii), \( (t, s, \alpha) \) for (iv).

Now let \( P(Z - \mu = -\sigma^2/a) = p \), \( P(Z - \mu = a) = 1 - p \). With \( a = s \) and \( a = t \), equality is attained in cases (v) and (vi), respectively. Equality in the case (vii) is attained if \( P(Z - \mu = s) = -\sigma^2 / [s(t - s)] \), \( P(Z - \mu = t) = \sigma^2 / [t(t - s)] \).

It should be noted that \( \sigma^2 \leq -\alpha \beta \) since \( \alpha \leq X - \mu \leq \beta \).

(This inequality, as well as others, has been obtained by Melvin Dresher and Albert Madansky using the theory of moment spaces.)

To obtain the corresponding enclosure Theorem 2.3, we examine the constraints which guarantee that the bound is non-negative. The intervals obtained for cases (i), (ii), (v) and (vi) are larger than those for cases (iii) and (iv). The latter two combined yield the result that there is an eigenvalue of \( H - \mu I \) in \([\alpha, r] \cup [-\sigma^2 / r, \beta]\) for \( \alpha \leq r \leq -\sigma^2 / \beta \), i.e., there is an eigenvalue of \( H \) in.
[α + μ, r + μ] ∪ [μ - σ^2 / r, β + μ]. The hypotheses α + μ = 0, β + μ = 1 give the desired result.

4. An Enclosure Theorem for Pairs of Eigenvalues.

The following theorem is due to Hájek and Rényi [8].

4.1 Theorem.

If X_1, ..., X_n are independent random variables, EX_j = 0, EX_j^2 = σ_j^2, 0 ≤ t_1σ_1 ≤ ... ≤ t_nσ_n, then

P( |X_1 + ... + X_m| ≤ t_mσ_m, m = 1, 2, ..., n ) ≥ 1 - ∑_j t_j^{-2}.

To obtain the corresponding enclosure theorem, define μ_i = x_iHx_i^*, σ_i^2 = x_i(H - μ_iI)^2x_i^*, i = 1, 2. Hence there is a pair of eigenvalues, not necessarily distinct, in the region μ_1 ± t_1σ_1, μ_1 + μ_2 ± t_2σ_2; for all 0 ≤ t_1σ_1 ≤ t_2σ_2, t_1^{-2} + t_2^{-2} ≤ 1.

Thus, for example, if t_1^2 = 1/λ, t_2^2 = 1 / (1 - λ), with σ_1^2 / σ_2^2 ≤ λ / (1 - λ), we obtain the region

μ_1 ± σ_1 / √λ, μ_1 + μ_2 ± σ_2 / √1-λ.

5. Remarks on Kantorovich's Inequality.

We now show how the probabilistic interpretation yields results concerning inequalities for quadratic forms.

If H is an Hermitian matrix such that 0 < m ≤ zHz^* ≤ M when z is a vector of unit length, Kantorovich's inequality [9] asserts

\[ P ( |X_1 + ... + X_m| ≤ t_mσ_m, m = 1, 2, ..., n ) ≥ 1 - ∑_j t_j^{-2}. \]
\[ z^* \prod_{z} z^{-1} \leq (m+M)^2/(4mm) \]

The probabilistic interpretation of this is that if \( X \) is a random variable with \( P(m \leq X \leq M) = 1 \) \( (m > 0) \), then

\begin{equation}
(5.1) \quad \frac{E}{EX^{-1}} \leq \frac{(m+M)^2}{(4mm)}.
\end{equation}

This inequality cannot be improved. However, if \( EX = \mu \) is known, then the inequality from (5.1)

\[ EX^{-1} \leq \frac{(m+M)^2}{(4mm\mu)} \]

is not sharp, but there is an improvement, namely,

\begin{equation}
(5.2) \quad EX^{-1} \leq \frac{(m+M-\mu)}{(m\mu)}.
\end{equation}

This inequality is based upon the simple fact that if \( f(x) \geq g(x) \) for \( x \in a \), then \( \int_{a} f(x) \, dx \geq \int_{a} g(x) \, dx \). In particular, if \( g(x) \) is convex on the interval \( (m, M) \) and \( f(x) \) is the chord through the points \( (m, g(m)) \) and \( (M, g(M)) \), we obtain

\begin{equation}
(5.3) \quad E g(x) \leq \frac{g(M) - g(m)}{M - m} \cdot EX + \frac{M \cdot g(m) - m \cdot g(M)}{M - m}.
\end{equation}

Equality is attained for the distribution \( P(X = m) = 1 - P(X = M) = (M - \mu) / (M - m) \).

Of interest are the special cases \( g(x) = x^{-1} \) and \( g(x) = x^k \), \( k > 1 \).
A number of proofs of (5.1) have been given, e.g., see Henrici [10]. We note that a simple proof is based upon (5.2), since

\[ \text{EX EX}^{-1} \leq \left( \frac{m+M}{mM} \right) \mu - \frac{\mu^2}{mM} \leq \frac{(m+M)^2}{4mM} . \]

Another form for (5.1) is the following: If \( P [m \leq X \leq M] = 1 \), \( m > 0 \), \( P (0 < z) = 1 \), \( E Z = 1 \), then

\[ (5.4) \quad \text{EX EX}^{-1} \leq \frac{(m+M)^2}{4mM} . \]

The proof is essentially the same as for (5.1). A direct application of this inequality is the case where \( A \) and \( B \) are permutable Hermitian matrices with \( 0 < m_1 \leq zAz^* \leq M_1 \), \( 0 < m_2 \leq zBz^* \leq M_2 \), \( zz^* = 1 \), whence

\[ zA^2z^* zB^2z^* \leq \frac{(m_1 m_2 + M_1 M_2)^2}{M_1 M_2 M_1 M_2} (zA^2z^*)^2 . \]
REFERENCES

transcendante de Fredholm les plus voisines d'une
nombre donné par les méthodes des moindres carrés


71 (1947), pp. 902-904.

426.

pp. 607-608.


Tchebycheffschen lemmas," Skand. Aktuarietidskr.,
23 (1940), pp. 121-125.

[8] J. Hájek and A. Rényi, "Generalization of an inequality of
6 (1955), pp. 281-283.