Approximation numbers

Seidel, J.J.

Published: 01/01/1979

Document Version
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
- A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 11. Dec. 2018
Approximation Numbers
Extended Notes of a Lecture by A.A. Melkman

by

J. J. Seidel

Technological University
Department of Mathematics
P.O. Box 513, Eindhoven
The Netherlands
Approximation numbers.


by J.J. Seidel

1. Summary.

The problem is to approximate, in the $p$-norm, the unit matrix $I_m$ of size $m$ by rank $n$ matrices, $n \leq m$, that is, to determine the numbers

$$ a_n = a_n (I_m; \ell_1^m, \ell_p^m) := \min_{A \in A_n} \max_{\|x\|_1 = 1} \| (I_m - A)x \|_p, $$

where $A_n$ denotes the set of all $m \times m$ matrices of rank $\leq n$.

Since only the extremes of the octahedron $\sum_{i=1}^{m} x_i \leq 1$ are of importance, this is equivalent to

$$ a_n = \min_{A \in A_n} \max_{1 \leq k \leq m} \| e_k - Ae_k \|_p, $$

where $e_1, \ldots, e_m$ denote the columns of $I_m$. In terms of subspaces $X_n \subset \mathbb{R}^m$, this is equivalent to

$$ a_n = \min_{X_n \subset \mathbb{R}^m} \max_{1 \leq k \leq m} \min_{x \in X_n} \| e_k - x \|_p. $$

It follows from a theorem by Sofman [5] that

$$ a_n (I_m; \ell_1^m, \ell_2^m) = \sqrt{1 - \frac{n}{m}}. $$

Melkman [1] proves that for $p \geq 2$

$$ a_n (I_m; \ell_1^m, \ell_p^m) \geq \left[ 1 + (m - 1) \left( \frac{(m - 1)(m - n)}{n} \right) - \frac{p}{2(p - 1)} \right]^{-1} + \frac{1}{p}. $$

For the case $p = \infty$ this amounts to
Moreover, Melkman shows that for \( p \neq 2 \) equality holds if and only if there exists a regular two-graph on \( m \) vertices with the multiplicities \( n \) and \( m - n \).

2. Eutactic stars ([4], [3])

Let \( X_m \) denote a real inner product space with an orthonormal basis \( e_1, \ldots, e_m \). Let \( P : X_m \to X_m \) be a projection operator (linear, symmetric, idempotent), and let \( PX_m = X_n \). Then, for \( k, l = 1, \ldots, m \),

\[
(e_k, Pe_l) = (Pe_k, Pe_l) = (Pe_k, e_l),
\]

hence for the vectors \( Pe_1, \ldots, Pe_m \in X_n \) the Gram matrix equals the coordinate matrix with respect to \( e_1, \ldots, e_m \). This matrix is symmetric and idempotent, has trace \( n \) hence

\[
\sum_{k=1}^{m} (Pe_k, Pe_k) = n.
\]

By definition, the vectors \( Pe_1, \ldots, Pe_m \) constitute a eutactic star. This star is spherical whenever all \( (Pe_k, Pe_k) \) are equal. If, in addition, \( |(Pe_k, Pe_{-l})| \) is constant for all \( k \neq l \), then the lines spanned by \( Pe_1, \ldots, Pe_m \) constitute a set of \( m \) equiangular lines in \( X_n \) at \( \cos \psi = \sqrt{\frac{m - n}{n(m - 1)}} \). Such a set is extremal in the following sense. For any set of \( m \) equiangular lines in \( \mathbb{R}^n \) at angle \( \psi \), let \( G \) denote the Gram matrix of a set of \( m \) unit vectors, one along each of the lines. Then \( G = I + C \cos \psi \), where \( C \) is a symmetric matrix of size \( m \) with diagonal zero and entries \( \pm 1 \) elsewhere. Since \( G \) is positive semidefinite of rank \( n \), its nonzero eigenvalues \( \lambda_1, \ldots, \lambda_n \) satisfy

\[
m = \text{tr} \ G = \lambda_1 + \ldots + \lambda_n , \quad m + m(m - 1) \cos^2 \psi = \text{tr} \ G^2 = \lambda_1^2 + \ldots + \lambda_n^2 ,
\]

hence

\[
\frac{m^2}{n} \leq m + m(m - 1) \cos^2 \psi , \quad \frac{m - n}{n(m - 1)} \leq \cos^2 \psi.
\]
Equality holds iff \( \lambda_1 = \ldots = \lambda_n \), that is, iff \( C \) has just two eigenvalues, of multiplicities \( n \) and \( m-n \).

A triple of equiangular lines is of acute or obtuse type, according as the lines are spanned by a triple of equiangular vectors at acute or at obtuse angle. An extremal set of \( m \) equiangular lines in \( \mathbb{R}^n \) is characterized by the property that each pair of lines is in a constant number of triples of obtuse type. This is equivalent to the existence of a regular two-graph on \( n \) vertices, whose eigenvalues have the multiplicities \( n \) and \( m-n \). For the definition and a survey cf. [3].

3. Sofman's theorem ([5], see also [1]).

Theorem. Let \( e_1, \ldots, e_m \) be an orthonormal basis of \( X_m \).

The conditions \( \sum_{k=1}^m \xi_k^2 = n \), \( 0 \leq \xi_k \leq 1 \) are necessary and sufficient for the existence of a subspace \( X_n \) of \( X_m \) such that, for \( k = 1, \ldots, m \),

the projection of \( e_k \) onto \( X_n \) has length \( \xi_k \).

Proof. The necessity of the condition has been observed in 2. For the sufficiency we use induction on \( n \).

Suppose, for \( k = 1, \ldots, m \),

\[
0 \leq \xi_k \leq 1, \quad \max_k \xi_k \leq \xi_1', \quad \xi_1'^2 + \xi_2'^2 + \ldots + \xi_m'^2 = n.
\]

Then \( \sum_{k=2}^m \xi_k^2 = n - \xi_1'^2 \geq n - 1 \), so we may choose \( \eta_1', \ldots, \eta_m \) such that \( \eta_1 = 0 \),

\[
0 \leq \eta_2 \leq \xi_2', \ldots, 0 \leq \eta_m \leq \xi_m', \quad \sum_{k=1}^m \eta_k^2 = n - 1.
\]

By the induction hypothesis there exists \( X_{n-1} \subset X_m \) such that, for \( k = 1, \ldots, m \),

the projection of \( e_k \) onto \( X_{n-1} \) has length \( \eta_k \). Then \( X_{n-1} \perp e_1 \). We define

\[
X_n := \langle e_1 \rangle \oplus X_{n-1}.
\]

Then, for \( k = 1, \ldots, m \), the projections \( z_k \) of \( e_k \) onto \( X_n \) have lengths

\[
\zeta_1 = 1, \quad \zeta_2 = \eta_2, \ldots, \zeta_m = \eta_m, \quad \sum_{k=1}^m \zeta_k^2 = n.
\]
We wish to rotate \( e_1, \ldots, e_m \) so as to move step by step from \( \xi_1, \ldots, \xi_m \) to \( \xi_1, \ldots, \xi_m \). Take any \( k = 2, \ldots, m \), and consider a rotation about \( \alpha \) in the plane \( \langle e_1, e_k \rangle \), leaving all other \( e_i \) fixed:

\[
\begin{align*}
e_1(\alpha) &= e_1 \cos \alpha - e_k \sin \alpha, \\
e_k(\alpha) &= e_1 \sin \alpha + e_k \cos \alpha, \\
e_1(\alpha) &= e_1.
\end{align*}
\]

Then \( \xi_j(\alpha) := \| P(e_j(\alpha)) \| \) satisfy \( \xi_1^2(\alpha) + \xi_k^2(\alpha) = \xi_1^2 + \xi_k^2 \), and

\[
\xi_i(\alpha) = \xi_i \quad \text{for} \quad i \neq 1, k.
\]

Since \( \xi_k(\alpha) \) is an increasing function of \( \alpha \) from \( \xi_k(0) = \xi_k \) to \( \xi_k(\frac{\pi}{2}) = \xi_1 = 1 \), and since \( 1 = \xi_1 \geq \xi_2 \geq \xi_k \), we may choose \( \alpha \) such that \( \xi_k(\alpha) = \xi_1 \). Then \( \xi_1(\alpha) = \xi_1 \leq \xi_1 \) implies

\[
\sum_{i=1}^{m} \xi_i^2(\alpha) = n, \quad \sum_{i=1}^{m} \xi_i^2 \quad \text{and} \quad \xi_1 \leq \xi_1 \leq \xi_1(\alpha).
\]

Thus we have made \( \xi_k \) into \( \xi_1 \), and \( \xi_1(\alpha) \) is again the largest number. Now repeat the process with each of the indices \( \neq 1 \), so as to arrive at an orthonormal basis whose projections onto \( X_n \) have given lengths \( \xi_2, \ldots, \xi_m \).

Then also the first length fits with \( \xi_1 \). This finishes the induction step from \( n - 1 \) to \( n \). For \( n = 1 \) the theorem is true, since then the line spanned by the vector \( (\xi_1, \ldots, \xi_m) \) applies. Thus the proof is completed.

The theorem may be rephrased in the following ways.

**Corollary.** Necessary and sufficient for the existence of a eutactic star in \( X_n \) consisting of \( m \) vectors at lengths \( \xi_1, \ldots, \xi_m \) are

\[
\xi_1^2 + \ldots + \xi_m^2 = n, \quad 0 \leq \xi_k \leq 1.
\]
Corollary. Necessary and sufficient for the existence of a symmetric idempotent matrix $C$ are:

$$\text{trace } C = \text{rank } C, \quad \text{diag } C > 0$$

In particular, taking $\xi_1 = \ldots = \xi_m = \sqrt{\frac{n}{m}}$ we have:

Corollary. $X_m$ contains spherical eutactic stars of any cardinality $m$.

Corollary. Given $m, n \in \mathbb{N}$, $m \geq n$, there exists a symmetric zero-diagonal matrix of size $m$ whose only eigenvalues are $n$ and $n - m$.

Corollary. $a_n(I_m, \hat{x}_1^m, \hat{x}_2^m) = \sqrt{1 - \frac{n}{m}}$.

Proof. For the $\ell_2$-norm we know

$$\min_{x \in X_n} \| e_k - x \| = \| e_k - Pe_k \|.$$  

Now

$$\| Pe_1 \|^2 + \ldots + \| Pe_m \|^2 = n,$$

$$\| e_1 - Pe_1 \|^2 + \ldots + \| e_m - Pe_m \|^2 = m - n$$

implies

$$\max_{1 \leq k \leq m} \| e_k - Pe_k \| > \frac{1 - \frac{n}{m}}{1 - \frac{n}{m}}.$$

By Sofman's theorem, an $X_n$ with $\| Pe_1 \| = \ldots = \| Pe_m \| = \sqrt{\frac{n}{m}}$ really exists, hence for such $X_n$:

$$\max_{1 \leq k \leq m} \| e_k - Pe_k \| = \sqrt{1 - \frac{n}{m}}.$$

Remark. Any $X_n$, for which

$$\max_{1 \leq k \leq m} \| e_k - Pe_k \| = \sqrt{1 - \frac{n}{m}},$$

satisfies $\| Pe_1 \| = \ldots = \| Pe_m \| = \sqrt{\frac{n}{m}}$. Indeed,
\[ \| P_{k} + P_{k} \|^2 + \| e_{k} - P_{k} \|^2 = 1, \]

\[ n + \sum_{k=1}^{m} \| e_{k} - P_{k} \|^2 = m = n + \sum_{1 \leq k \leq m} \| e_{k} - P_{k} \|^2, \]

\[ \forall 1 \leq k \leq m \| e_{k} - P_{k} \|^2 = \max_{1 \leq k \leq m} \| e_{k} - P_{k} \|^2, \]

hence all \( \| P_{k} \| \) are equal.

4. The theorem of Hahn - Banach.

Let \( V \) be a Banach space with norm \( \| \cdot \| \), and let \( W \) be a closed subspace. The quotient space \( V/W \) is a Banach space with norm

\[ \| v + w \| := \inf_{w \in W} \| v - w \|. \]

The dual space \( V^* \) of all continuous linear functionals \( f \) on \( V \) is a Banach space with norm

\[ \| f \|_* := \sup_{\| x \| = 1} \| f(x) \| = \sup_{0 \neq x \in V} \frac{|f(x)|}{\| x \|}. \]

The following is a consequence of the Hahn - Banach theorem.

Theorem. \[ \| x \| = \sup_{\| f \|_* = 1} \| f(x) \| = \sup_{0 \neq f \in V^*} \frac{|f(x)|}{\| f \|_*}. \]

Furthermore, we need the following

Theorem. There exists an isometric isomorphism of the spaces \( (V/W)^* \) and \( W^\perp := \{ f \in V^* \mid f(W) = 0 \} \).

We apply the above to the linear space \( X_m \) provided with the \( l_p \)-norm

\[ \| x \|_p := \left( \sum_{i=1}^{m} |x_i|^p \right)^{1/p}. \]
Hölder: \[ \frac{1}{p} \sum_{i=1}^{m} |x_i y_i| \leq \left( \frac{1}{p'} \sum_{i=1}^{m} |x_i|^p \right)^{1/p} \left( \frac{1}{p'} \sum_{i=1}^{m} |y_i'|^{p'} \right)^{1/p'} \]

with \( \frac{1}{p} + \frac{1}{p'} = 1, 1 < p < \infty \).

As a consequence, \( X^*_n \) has \( \ell_p \)-norm

\[ \|y\|_p = \left( \frac{1}{p} \sum_{i=1}^{m} |y_i|^p \right)^{1/p}, \quad \frac{1}{p} + \frac{1}{p'} = 1, \quad y \in X^*_n. \]

For a subspace \( X \) of \( X_n \) we now have

\[ X^*_n = \{ y \in X^*_m \mid y_1 x_1 + \ldots + y_m x_m = 0, \quad \forall x \in X_n \}, \]

where \( y_i = y(e_i) \). Apply Hahn-Banach to the quotient \( V/W = X/X_n \), then we obtain (since the dimension is finite):

**Theorem.** \( \| e_k + x \|_p = \inf_{x \in X_n} \| e_k - x \|_p = \max_{0 \neq y \in X^*_n} \| y \|_p \). \]

In 6. we will use the following consequence of Hölder's inequality.

**Theorem.** \[ \left( \sum_{j=1}^{m} |x_j|^p \right)^{1/p} \geq \left( \sum_{j=1}^{m} |x_j|^r \right)^{1/r} (m - 1)^{1 - \frac{1}{r}}, \quad \text{for } p > r. \]

Indeed, apply Hölder with \( q \geq 1 \) to the \( (m - 1) \)-vectors

\[ |x_1|^r, \ldots, |x_k|^r, \ldots, |x_m|^r \text{ and } 1, \ldots, 1, \text{ then} \]

\[ \left( \sum_{j \neq k}^{m} |x_j|^r \right)^{1/r} (m - 1)^{1 - \frac{1}{r}} \leq \left( \sum_{j \neq k}^{m} |x_j|^q \right)^{1/q} \frac{1}{r} \]

For \( p = r \), that is \( q = 1 \), our inequality is an equality.

For \( p > r \), equality holds iff the \( (m - 1) \)-vectors are proportional.
5. Melkman's theorem for \( p = \infty \).

In the following proof we will use back and forth the consequence of Hahn-Banach exposed in 4. In addition, we will use the Cauchy-Schwarz inequality in the following form:

\[
\sum_{i=1}^{m} |y_i| \leq \sqrt{\sum_{i=1}^{m} |y_i|^2} \sqrt{m - 1}.
\]

We first prove the \( p = \infty \) case, since this case is representative.

**Theorem.** \( a_n \left( I_m, x_1^m, x_\infty^m \right) \geq \left( 1 + \sqrt{(m-1)n \over m-n} \right)^{-1} \),

and equality holds iff there exists a regular two-graph on \( m \) vertices with multiplicities \( n, m-n \).

**Proof.** For \( k \in \{1, 2, \ldots, m\} \) we have

\[
\min_{x \in x_n} \|e_k - x\|_\infty = \max_{0 \neq y \perp x_n} \|Y\|_1
\]

\[
= \max_{0 \neq y \perp x_n} \frac{|y_k|}{|y_k| + \sum_{i \neq k} |y_i|} \geq \max_{0 \neq y \perp x_n} \frac{|y_k|}{|y_k| + \sum_{i=1}^{m} |y_i|^2 - |y_k|^2} = \max_{0 \neq y \perp x_n} \left[ 1 + \sqrt{m-1} \sqrt{\sum_{i=1}^{m} |y_i|^2} \right]^{-1} = \left[ 1 + \sqrt{m-1} \sqrt{\max_{0 \neq y \perp x_n} \left( \frac{|y_k|}{\|y\|_2} \right)^{-2} - 1} \right]^{-1}
\]
\[
\left[ 1 + \sqrt{m-1} \sqrt{\left( \min_{x \in X_n} \| e_k - x \|_2 \right)^{-2} - 1} \right]^{-1}.
\]

Since \( f(z) := \left[ 1 + \sqrt{m-1} \sqrt{z^{-2} - 1} \right]^{-1} \) is a monotone increasing function of \( z \), this implies

\[
\min_{X_n \subset X_m} \max_{1 \leq k \leq m} \min_{x \in X_n} \| e_k - x \|_\infty > \left[ 1 + \sqrt{m-1} \sqrt{\left( \min_{X_n \subset X_m} \max_{1 \leq k \leq m} \min_{x \in X_n} \| e_k - x \|_2 \right)^{-2} - 1} \right]^{-1}.
\]

Thus we have expressed \( a_n(I_m; \ell^m_1, \ell^m_\infty) \) in terms of \( a_n(I_m; \ell^m_1, \ell^m_2) \), which equals \( \sqrt{1 - \frac{n}{m}} \) by 3. Substitution yields the inequality of the theorem.

Now suppose we have equality:

\[
a_n(I_m; \ell^m_1, \ell^m_\infty) = \left[ 1 + \sqrt{\frac{n(m-1)}{m-n}} \right]^{-1}.
\]

We analyse the various steps performed in the proof of the inequality. First, for an optimal \( X_n \) we have

\[
\max_{1 \leq k \leq m} \min_{x \in X_n} \| e_k - x \|_2 = \sqrt{1 - \frac{n}{m}}
\]

hence, by the remark at the end of 3,

\[
\min_{x \in X_n} \| e_k - x \|_2 = \sqrt{1 - \frac{n}{m}}, \text{ for all } k = 1, \ldots, m.
\]

In addition, for any \( 1 \leq k \leq m \), the \( x \) which achieves \( \min_{x \in X_n} \| e_k - x \|_\infty \), also minimizes \( \| e_k - x \|_2 \). Hence the matrix \( A \) which approximates \( I_m \) in the \( \ell_\infty \)-norm also works for the \( \ell_2 \)-norm.
Secondly, if equality holds in Cauchy-Schwarz, then the corresponding vectors of length \( m - 1 \):

\[
( |Y_1|, |Y_2|, \ldots, \sqrt{\frac{\sum_{k=1}^{m} Y_k^2}{n}} \ldots, |Y_m| ) \text{ and } (1,1,\ldots,1)
\]

are proportional. So we may take \( |Y_1| = \ldots = |Y_m| = 1 \) except for \( |Y_k| \) which equals

\[
|Y_k| = \sqrt{\frac{(m-1)(m-n)}{n}}, \text{ since } \frac{|Y_k|}{|Y_k| + m-1} = \frac{1}{1 + \sqrt{\frac{n(m-1)}{m-n}}}
\]

For each \( k = 1, \ldots, m \) we find such a vector \( y_k \), which is proportional to the projection of \( e_k \) onto \( X_n^1 \). These vectors are taken as the columns of the following coordinate matrix \( B \) ( = Gram matrix, cf. 2) of rank \( m - n \):

\[
B = \begin{bmatrix}
\varepsilon_{11} & \varepsilon_{12} & \ldots & \varepsilon_{1m} \\
\varepsilon_{21} & \varepsilon_{12} & \ldots & \varepsilon_{2m} \\
\varepsilon_{m1} & \varepsilon_{m2} & \ldots & \varepsilon_{m1} \\
\end{bmatrix}
\]

It follows that the lines spanned by these vectors are equiangular, at

\[
\cos^2 \chi = \frac{n}{(m-1)(m-n)}.\]

This set of equiangular lines in \( X_n^1 \) is extremal in the sense of 2 (just interchange \( n \) and \( m - n \)). Equivalently, \( X_n \) contains an extremal set of equiangular lines at \( \cos^2 \varphi = \frac{m-n}{(m-1)n} \). Equivalently, there exists a regular two-graph on \( m \) vertices with the multiplicities \( n \) and \( m - n \).

6. Melkman's theorem for \( p > 2 \).

For \( p \geq r \), a lower bound for \( a_n(I_m ; \chi^m_1, \chi^m_2) \) in terms of \( a_n(I_m ; \chi^m_1, \chi^m_2) \) is obtained, and specialized to the case \( r = 2 \), since

\[
a_n(I_m ; \chi^m_1, \chi^m_2) = \sqrt{1 - \frac{n}{m}} \frac{n}{m} \text{ is known. Instead of Cauchy-Schwarz we use the following consequence of Hölder, cf. 4:}
\]
for \( p > r \); for \( p = r \) this degenerates into an equality.

For a fixed \( k \), we abbreviate as follows:

\[
\eta_p := \min_{x \in X} \| e_k - x \| = \max_{0 \neq y \perp X} \frac{|y_k|}{\|y\|^p}.
\]

Then for \( p \geq r \),

\[
\left( \eta_p^p \right)^{\frac{1}{p'}} = \max_{y \perp X} \frac{|y_k|^{p'}}{|y_k|^{p'} + \sum_{j \neq k} |y_j|^{p'}} \geq \max_{y \perp X} \frac{|y_k|^{p'}}{|y_k|^{p'} + (m - 1) 1 - \frac{1}{r'} \left( \sum_{j \neq k} |y_j|^{r'} \right)^{\frac{p'}{r'}}} = \max_{y \perp X} \left[ 1 + (m - 1) 1 - \frac{p'}{r'} \left( \max_{y \perp X} \frac{|y_k|}{\|y\|^r} \right)^{\frac{p'}{r'}} \right]^{-1} \]

It follows that

\[
\eta_p \geq \left[ 1 + (m - 1) 1 - \frac{p'}{r'} \left( \eta_r^{-r'} - 1 \right) \right]^{-\frac{1}{p'}}.
\]

**Theorem.** For \( p > 2 \)

\[
a_n(\mathbb{I}_m; z_1^m, z_p^m) \geq \left| 1 + (m - 1) \left( \frac{(m - 1)(m - n)}{n} \right)^{-\frac{p}{2(p - 1)}} \right|^{-1 + \frac{1}{p}},
\]
and equality holds iff there exists a regular two-graph on \( m \) vertices with multiplicities \( n, m - n \).

Proof. Put \( r = 2 \) in the formula for \( n_p \), proceed as in 5, and substitute the value of \( a_n \) for \( r = 2 \). This yields the inequality. For the case of equality we must have equality in the consequence of Hölder's inequality. Since \( p > 2 \), again the \((m - 1)\)-vectors are proportional, and the reasoning of 5 works.

Remark. The second part of the reasoning above does not work for \( p = 2 \). Indeed, then the consequence of Hölder's inequality is an equality, and yields nothing new. In fact, for \( p = 2 \) we do have

\[
a_n(\mathbb{1}^m; x_1^m, x_2^m) = \sqrt{1 - \frac{n}{m}},
\]

but the extremal sets need not be extremal sets of equiangular lines. Any spherical eutactic star provides an extremal set. As an example we mention the root systems [4].

[1] A.A. Melkman, The distance of a subspace of \( \mathbb{R}^m \) from its axes and \( n \)-widths of octahedra, to be published.