

How bad are symmetric Pick matrices*

Vadim Olshevsky
Department of Mathematics
Georgia State University
Atlanta, GA 30303
volshevsky@cs.gsu.edu
<http://www.cs.gsu.edu/~matvro>

Dario Fasino
Dipartimento di Matematica e Informatica
University of Udine, Udine, Italy

Abstract

Let P be a symmetric positive definite Pick matrix of order n . The following facts will be proven here: 1. P is the Gram matrix of a set of rational functions, with respect to a inner product defined in terms of a “generating function” associated to P ; 2. Its condition number is lower-bounded by a function growing exponentially in n . 3. P can be effectively preconditioned by the Pick matrix generated by the same nodes and a constant function.

1 Introduction

We consider here the class of real symmetric Pick matrices $P \equiv (p_{ij})$ defined as

$$p_{ij} = \frac{d_i + d_j}{x_i + x_j}, \quad (1)$$

for $i, j = 1 \dots n$, where the numbers $x_1 \dots x_n$ are pairwise distinct and positive. For notational simplicity, the nodes are supposed to be ordered, $x_1 > x_2 > \dots > x_n$. The latter hypothesis is of no restriction, since it can be fulfilled by a suitable row and column permutation on the original matrix. This class is among the most notable examples of matrices possessing a displacement structure [9]: In fact, any matrix P as in (1) is the solution of a Sylvester equation having the form $D_x P + P D_x = d e^* + e d^*$, where $e = (1 \dots 1)^*$, $d \equiv (d_i)$ and $D_x = \text{Diag}(x_1 \dots x_n)$ is the diagonal matrix with x_i as i th diagonal entry. Matrices with definitions slightly different from (1) are also found in the literature, which bear the same name, for example

$$\hat{P} \equiv \left(\frac{1 - c_i c_j}{x_i + x_j} \right),$$

where $-1 < c_i < 1$. These two definitions (and also other) are related by a diagonal congruence: Indeed, let $\Delta = \text{Diag}(1 + c_1 \dots 1 + c_n)$. Then $P = \Delta^{-1} \hat{P} \Delta^{-1}/2$ is the matrix (1) with $d_i = (1 - c_i)/(1 + c_i)$.

*This work was supported by NSF grant CCR 9732355.

Pick matrices arise in the classical Nevanlinna-Pick interpolation problem [2, 5]. In its simplest form, that problem amounts to find an analytic function mapping the unit disk into the closed unit disk which satisfies a set of prescribed interpolation conditions. Variants of this problem are also considered, where the unit disk is replaced by a half plane, and these variants make up the differences among the various definitions of Pick matrix. The solvability criterion for the Nevanlinna-Pick problem in all its forms is the positive-semidefiniteness of a Pick matrix defined in terms of the given interpolation points.

Interest in the Nevanlinna-Pick interpolation problem and its confluent or degenerate cases [5, 13], as well as its generalizations to matrix- and operator-valued functions [12], and the the quest for practical solution algorithms [10], has become particularly intense in the recent years, due to its deep connections with problems in time-invariant system theory, particularly, in prediction theory and control theory [2]. Moreover, Pick matrices occur in the Pták-Young generalization of the Schur-Cohn theorem [1].

This paper focuses on the spectral conditioning of symmetric positive definite Pick matrices. The message to be gained is that, all matrices in this class are very ill-conditioned, hence checking their positive definiteness is a numerically difficult task. Indeed, it is well known that a very large condition number can spoil the positive definiteness in finite precision arithmetic.

In the next section, a “generating function” will be associated to any given Pick matrix. This function make it possible to consider a Pick matrix as the Gram matrix of a set of rational functions, with respect to a suitable inner product. This Gram structure is at the basis of the results in Section 3 and 4. There, firstly we derive a lower bound for the condition number of P that depends exponentially on the order n . Furthermore, it will be shown that particular Pick matrices, whose generating function is constant, turn out to be good preconditioners for all matrices in the class considered here.

The following notations are used throughout this paper: All matrices are n -by- n . Let $\lambda_{\max}(P) = \lambda_1(P) \geq \dots \geq \lambda_n(P) = \lambda_{\min}(P)$ be the eigenvalues of P , and $\kappa_2(P)$ its spectral condition number. The purely imaginary numbers $\xi_i = ix_i$ will be considered besides the real nodes x_i . Moreover, let

$$\pi(x) = \prod_{i=1}^n (x - \xi_i).$$

Finally, observe that if $p(x)$ is a polynomial of degree less than n then

$$\frac{p(x)}{\pi(x)} = \sum_{i=1}^n \frac{v_i}{x - \xi_i}, \quad v_i = \frac{p(z_i)}{\pi'(\xi_i)}. \quad (2)$$

2 Generating functions for Pick matrices

Let $0 < a \leq +\infty$ be fixed, and let $w(x)$ be any real, even function, defined in the *reference interval* $(-a, a)$, such that

$$\int_{-a}^a \frac{1}{x^2 + x_i^2} w(x) dx = \frac{d_i}{x_i}, \quad i = 1 \dots n. \quad (3)$$

Hereafter, we call any such function w a *generating function* for the matrix P . Its existence is not a problem, for if we choose a set of linearly independent real functions $w_1(x) \dots w_n(x)$, with the property of being defined on $(-a, a)$ and even, then we can compute a function $w(x) = \sum_{i=1}^n c_i w_i(x)$ fulfilling

conditions (3), by solving a linear system whose matrix is nonsingular. However, much more can be said. In what follows, a generating function $w(x)$ is said to be *good* if satisfies the constraints (3) with $a = +\infty$, and moreover there exist positive constants m, M such that, for all $x \in \mathbf{R}$, $m \leq w(x) \leq M < +\infty$. The next theorem states precisely that every positive definite Pick matrix admits a good generating function.

Theorem 2.1 *If P is positive definite, there exists a generating function for P , relative to the reference interval \mathbf{R} , which is bounded from above and below by positive constants. If P is positive semidefinite, and has rank k , then there exists a unique Borel measure generating P , whose spectrum consists precisely of k distinct points.*

The proof of the above theorem will be postponed to the end of this section, since it will be derived as a consequence of Theorem 2.3, which has an independent interest. To keep notation simple and coherent through the paper, $w(x) dx$ sometimes will be used instead of the more appropriate $d\mu(x)$, where $\mu(x)$ a regular Borel measure.

As an example of the singular case mentioned in Theorem 2.1, consider that the Pick matrix generated by a unit mass in zero is $P \equiv (x_i x_j)$, which clearly has rank one.

A more relevant example, which will play a special role in what follows, is the set of matrices generated by a constant function in \mathbf{R} , $w(x) \equiv c$. Then, in this case, $d_i = c\pi$, independently on the nodes x_i . We call such matrices *Hilbert-Pick matrices*. Indeed, one particular case in this class is the well known *Hilbert matrix* $H \equiv (1/(i + j - 1))$, which is generated by the function $w(x) \equiv 1/2\pi$ and the nodes $x_i = (2i - 1)/2$. We will show in Section 4 that all Hilbert-Pick matrices are spectrally close to all Pick matrices with same nodes.

Remark that a Hilbert-Pick matrix P satisfies the Lyapunov equation $D_x P + P D_x = c e e^*$ for some constant $c > 0$, hence it is positive definite, by Lyapunov theorem. Spectral properties of matrices defined as solutions of Lyapunov equations with a low rank right side are studied in [11], where it is shown that their eigenvalues have a fast decay. Moreover, Hilbert-Pick matrices are computationally very convenient: Since the coefficient matrices defining the above Lyapunov equation are diagonal, and the right side of that equation has rank one, there are stable and efficient algorithms for computing their triangular factors, see [8].

2.1 Pick matrices are Gram matrices

Theorem 2.2 *Consider the set of rational functions*

$$r_i(x) = \frac{1}{x - \xi_i} \quad i = 1 \dots n, \quad (4)$$

where $\xi_i = ix_i$. Then, the matrix P is the Gram matrix of the functions $r_i(x)$ with respect to the inner product

$$\langle f, g \rangle_w = \int_{-a}^a f(x) \bar{g}(x) w(x) dx,$$

i.e., $p_{ij} = \langle r_i, r_j \rangle_w$, where w is any generating function of P .

PROOF. In order to prove the equality $p_{ij} = \langle r_i, r_j \rangle_w$, consider separately the real part and the imaginary part, $\langle r_i, r_j \rangle_w = \Re \langle r_i, r_j \rangle_w + i \Im \langle r_i, r_j \rangle_w$. For the real part, we have

$$\begin{aligned}
\Re \langle r_i, r_j \rangle_w &= \frac{1}{2} (\langle r_i, r_j \rangle_w + \langle r_j, r_i \rangle_w) \\
&= \frac{1}{2} \int_{-a}^a \left(\frac{1}{(x - \xi_i)(x + \xi_j)} + \frac{1}{(x + \xi_i)(x - \xi_j)} \right) w(x) dx \\
&= \int_{-a}^a \left(\frac{1}{x^2 + x_i^2} \frac{x_i}{x_i + x_j} + \frac{1}{x^2 + x_j^2} \frac{x_j}{x_i + x_j} \right) w(x) dx \\
&= \frac{x_i}{x_i + x_j} \frac{d_i}{x_i} + \frac{x_j}{x_i + x_j} \frac{d_j}{x_j} \\
&= \frac{d_i + d_j}{x_i + x_j}.
\end{aligned}$$

The imaginary part vanishes,

$$\begin{aligned}
2i \Im \langle r_i, r_j \rangle_w &= \langle r_i, r_j \rangle_w - \langle r_j, r_i \rangle_w \\
&= \int_{-a}^a \left(\frac{1}{(x - \xi_i)(x + \xi_j)} - \frac{1}{(x + \xi_i)(x - \xi_j)} \right) w(x) dx \\
&= 0,
\end{aligned}$$

since the last integrand is odd. ■

Observe that a result similar to the one in the preceding theorem, but for a different definition of Pick matrix, can be found in [13], where the range of integration is the unit circle and the basis functions are polynomials.

As an immediate consequence, if P admits a nonnegative generating function, then P is positive definite. Now, in order to make a step toward the opposite conclusion, let $\omega(x) = \prod_{i=1}^n (x^2 + x_i^2)$, and define the inner product

$$\langle p, q \rangle_{\hat{w}} = \int_{-a}^a p(x) \bar{q}(x) \hat{w}(x) dx, \quad \hat{w}(x) = \frac{w(x)}{\omega(x)}. \quad (5)$$

That inner product is well defined whenever p, q are polynomials of degree less than n . Indeed, if $f(x) = p(x)/\pi_\xi(x)$ and $g(x) = q(x)/\pi_\xi(x)$, then from (2) we see that $f(x)$ and $g(x)$ are in the linear span of the functions defined in (4). Moreover, $\langle f, g \rangle_w = \langle p, q \rangle_{\hat{w}}$, since $\omega(x) = \pi_\xi(x) \pi_{-\xi}(x)$. In particular, let $p(x) = p_0 + \dots + p_{n-1} x^{n-1}$, and denote by p and v the n -vectors with entries p_i and $v_i = p(\xi_i) \pi'_\xi(\xi_i)$ respectively. Then from (2) follows:

$$\begin{aligned}
v^* P v &= \int_{-a}^a p(x) \bar{p}(x) \hat{w}(x) dx \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} p_i \bar{p}_j \int_{-a}^a x^{i+j} \hat{w}(x) dx \\
&= p^* H p,
\end{aligned} \quad (6)$$

where H is a Hankel matrix whose entries are the algebraic moments of $\hat{w}(x)$. This fact helps us to prove the following fact:

Theorem 2.3 *Let $V \equiv (\xi_i^{j-1})$ be the Vandermonde matrix with nodes ξ_i , and D be the diagonal matrix with entries $\pi'_\xi(\xi_1) \dots \pi'_\xi(\xi_n)$. Then*

$$V^* D^* P D V = H, \quad H \equiv \left(\int_{-a}^a x^{i+j} \hat{w}(x) dx \right).$$

PROOF. On the basis of (6), it suffices to observe that $DVp = v$. ■

2.2 Proof of Theorem 2.1

Let P be positive definite, and let $H \equiv (h_{i-j+2})$ be the Hankel matrix given by Theorem 2.3. Since H is congruent to P , it is positive definite, hence by classical results on the finite Hamburger moment problem [6, 14] there exists a nonnegative function $\hat{w}(x) \in L_1(\mathbf{R})$ such that

$$\int_{\mathbf{R}} x^i \hat{w}(x) dx = h_i, \quad i = 0 \dots 2n - 2. \quad (7)$$

A close look at the definition of H reveals that $h_{2k-1} = 0$, hence we can safely suppose that $\hat{w}(x)$ is even. Now, we see from (5) that $w(x) = \omega(x)\hat{w}(x)$ fulfills the equalities

$$\int_{\mathbf{R}} \frac{1}{x^2 + x_i^2} w(x) dx = \mu_i, \quad i = 1 \dots n,$$

where $\mu_i = d_i/x_i$. We have, equivalently,

$$\int_{\mathbf{R}} a_i(x) \frac{w(x)}{x^2 + 1} dx = \mu_i, \quad a_i(x) = \frac{x^2 + 1}{x^2 + x_i^2} \in L_\infty(\mathbf{R}).$$

Moreover, $(\mathbf{R}, (x^2 + 1)^{-1} dx)$ is a finite measure space and, by construction, $w(x) \in L_1(\mathbf{R}, (x^2 + 1)^{-1} dx)$. Hence, we are in the hypotheses of Theorem 2.9 of [3]. The consequence of that theorem is that we can assume $m \leq w(x) \in L_\infty(\mathbf{R})$, for some $m > 0$. The first part of the claim follows.

If P is semidefinite and has rank k , then the same holds also for H , hence the constraints (7) individuate a unique Borel measure $\hat{w}(x) dx$ whose spectrum consists of precisely k points, see [14, Thm. 1.2]. The conclusion follows again from the relation $w(x) = \omega(x)\hat{w}(x)$. ■

The preceding discussion gives us the basis for a procedure to compute a good generating function of a given P : Indeed, consider the convex functional

$$E(f) = \int_{\mathbf{R}} f(x) [\log f(x) - 1] \frac{dx}{x^2 + 1},$$

which is a “nonautonomous” variant of the Shannon entropy [3, 6]. Observe that $E(f)$ is finite for every nonnegative $f \in L_\infty(\mathbf{R})$. Hence, we may consider the minimization of $E(f)$ among all bounded functions which are nonnegative and fulfil the constraints (3). The minimizer exists, since the admissible set is closed and not empty, is unique, since $E(f)$ is strictly convex, and can be expressed by Lagrange multiplier theorem as

$$w(x) = \exp \left(\sum_{i=1}^n \lambda_i \frac{x^2 + 1}{x^2 + x_i^2} \right),$$

see [3, Thm. 4.8]. This function is, obviously, a good generating function for P . A similar argument, based on the minimization of the variant of the Burg entropy [3] given by

$$F(f) = - \int_{\mathbf{R}} \log f(x) \frac{dx}{x^2 + 1},$$

allows us to derive the existence of a good generating function having the rational form

$$w(x) = \left(\sum_{i=1}^n \lambda_i \frac{x^2 + 1}{x^2 + x_i^2} \right)^{-1}.$$

3 Exponential ill-conditioning

The purpose of this section is to prove that, for any symmetric positive definite Pick matrix P , $\kappa_2(P)$ is bounded from below by an exponential function in n , regardless of the generating function. This result is quite similar in spirit to other results available for Vandermonde [7], Krylov and Hankel matrices [15, 16], which are other examples of displacement structured matrices.

The starting point is a very simple technique: One chooses suitable vectors v, w and gets immediately the lower bound

$$\kappa_2(P) \geq \frac{v^* P v}{v^* v} \frac{w^* w}{w^* P w}.$$

In our case, the most convenient choice for v is a canonical basis vector, for example, $v = (1, 0 \dots 0)^*$, so that $v^* v = 1$ and

$$v^* P v = \int_{-a}^a \frac{1}{x^2 + x_1^2} w(x) dx.$$

To proceed further, consider the vector w whose entries are given by

$$\begin{aligned} w_1 &= \prod_{j \neq 1} \frac{\xi_1 + \xi_j}{\xi_1 - \xi_j} \\ w_i &= \frac{2\xi_i}{\xi_i - \xi_1} \prod_{j \neq 1, i} \frac{\xi_i + \xi_j}{\xi_i - \xi_j} \quad i = 2 \dots n. \end{aligned}$$

From (2) we have

$$\sum_{i=1}^n \frac{w_i}{x - \xi_i} = \frac{1}{\pi_\xi(x)} \prod_{i=2}^n (x + \xi_i),$$

hence

$$w^* P w = \int_{-a}^a \frac{1}{x^2 + x_1^2} w(x) dx = v^* P v.$$

So far, our estimate consists in the inequality $\kappa_2(P) \geq w^* w$ which, remarkably, does no longer depend on the generating function. It remains to estimate from below the squared norm of w . A rather crude bound is

$$w^* w > |w_1|^2 > \frac{(x_1 + x_n)^{2n-2}}{(x_1 - x_n)^{2n-2}}.$$

We obtain:

Theorem 3.1 *Let P be given by (1) and positive definite, and let $r = (x_1 + x_n)^2 / (x_1 - x_n)^2 > 1$. Then $\kappa_2(P) > r^{n-1}$.*

The conclusion in the above result looks weak, if compared to the above mentioned results concerning Vandermonde and positive definite Hankel matrices, since r can become arbitrarily close to 1. In fact, the condition numbers of the latter matrices can be bounded from below by exponentials like 2^n and 4^n respectively, apart of slowly varying factors, with no further assumptions.

Certainly, the preceding expression of r deserves some improvement. Nevertheless, the above theorem can be regarded as qualitatively optimal. Indeed, in Figure 1 we show some statistics concerning optimally conditioned Hilbert-Pick matrices. For $n = 2 \dots 20$, we have computed the minimum of the function $\kappa_2(P_n)$, where P_n is the Hilbert-Pick matrix with nodes $x_1 \dots x_n$, assumed as free parameters, under the constraints $x_i > 0$.

In Figure 1a, the computed minimum value of $\kappa_2(P_n)$ is plotted against n . Figures 1b and 1c show the plots of $\kappa_2(P_n)/\kappa_2(P_{n-1})$ and $(\kappa_2(P_n))^{1/n}$ respectively. These quantities are asymptotically equal, hence these numerical examples indicate that $(\kappa_2(P_n))^{1/n}$ becomes close to one, if the nodes are unbounded and chosen in an optimal way. For completeness, we show in Figure 1d the optimal nodes, normalized so that the least node is equal to 1.

Moreover, the conclusion of the above theorem is in agreement with the main result of [11], which shows that the bound for the eigenvalue decay of Hilbert-Pick matrices may become arbitrarily close to 1, when the ratio x_1/x_n diverges.

4 A preconditioning technique

The main result in this section is Corollary 4.3, stating that the Hilbert-Pick matrices introduced in Section 2 are good preconditioners for Pick matrices, since the conditioning of the preconditioned matrices does not depend on n . This result is the Pick counterpart of a similar result obtained in [4] for Hankel matrices.

The statement in the following lemma is a basic result in linear algebra, whose simple proof is omitted:

Lemma 4.1 *Let P_1, P_2 be hermitian positive definite matrices, and suppose that for all vectors v it holds*

$$m \leq \frac{v^* P_1 v}{v^* P_2 v} \leq M.$$

Then, $\lambda_i(P_2^{-1} P_1) \in [m, M]$, and

$$\kappa_2(P_2^{-1} P_1) \leq \frac{M}{m}.$$

Theorem 4.2 *Let P_1, P_2 be Pick matrices generated by nodes $x_1 \dots x_n$ and functions w_1, w_2 in the same reference interval $(-a, a)$, and suppose that for $|x| \leq a$ it holds*

$$m \leq \frac{w_1(x)}{w_2(x)} \leq M.$$

Then $\lambda_i(P_2^{-1} P_1) \in [m, M]$ and $\kappa_2(P_2^{-1} P_1) \leq M/m$, with equality if and only if $m = M$.

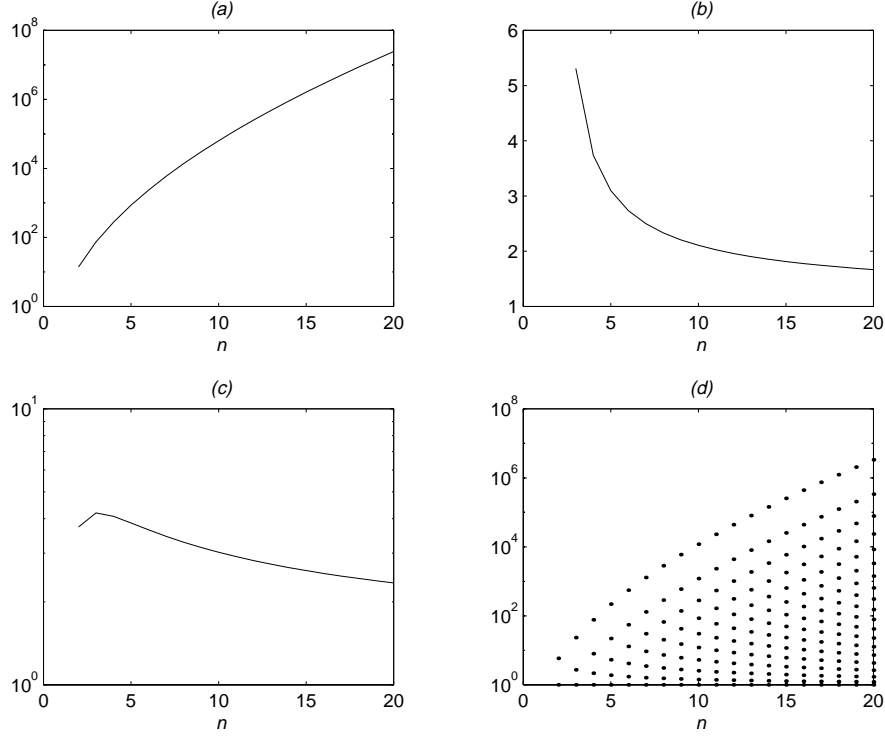


Figure 1: Optimally conditioned Hilbert-Pick matrices

PROOF. Let $\omega(x) = \pi_\xi(x)\pi_{-\xi}(x)$. For $v \equiv (v_i)$ let $p(x)$ be the polynomial defined by the interpolation conditions $p(\xi_i) = v_i\pi'_\xi(\xi_i)$. From (2) and Theorem 2.2 we have

$$mv^*P_2v = m \int_{-a}^a \frac{|p(x)|^2}{\omega(x)} w_2(x) dx \leq \int_{-a}^a \frac{|p(x)|^2}{\omega(x)} w_1(x) dx = v^*P_1v,$$

since all integrands are nonnegative. Analogously, one proves

$$v^*P_1v = \int_{-a}^a \frac{|p(x)|^2}{\omega(x)} w_1(x) dx \leq M \int_{-a}^a \frac{|p(x)|^2}{\omega(x)} w_2(x) dx = Mv^*P_2v.$$

Thesis follows from Lemma 4.1. ■

Corollary 4.3 *If $w(x)$ is a good generating function for P , $m \leq w(x) \leq M$ for $x \in \mathbf{R}$, and K is a Hilbert-Pick matrix with the same nodes as P , then $\lambda_i(K^{-1}P) \in [m, M]$ and $\kappa_2(K^{-1}P) \leq M/m$.*

Finally, we would like to gain some more insight about the distribution of the eigenvalues of $K^{-1}P$. This is the subject of the following proposition and the subsequent example.

Theorem 4.4 *Introduce the set*

$$\mathcal{P}_n = \left\{ f(x) = \sum_{i=1}^n v_i r_i(x), \quad \int_{\mathbf{R}} |f(x)|^2 dx = 1 \right\},$$

where the functions $r_i(x)$ are defined in (4). In the notations of Corollary 4.3 we have

$$\lambda_{\min}(K^{-1}P) = \min_{f \in \mathcal{P}_n} \int_{\mathbf{R}} |f(x)|^2 w(x) dx,$$

and analogously for the maximum.

PROOF. Let $L \equiv (l_{ij})$ be the inverse of the lower triangular Cholesky factor of K , and let

$$s_i(x) = \sum_{j=1}^n l_{ij} \frac{1}{x - \xi_i}.$$

From the relation $LKL^* = I$, the identity matrix, and Theorem 2.2, we see that the functions $s_i(x)$ are orthonormal with respect to the inner product $\langle \cdot, \cdot \rangle_w$. Indeed, the above construction of $s_i(x)$ is equivalent to applying the Gram-Schmidt process to the set of rational functions $r_i(x)$. Now, for any vector $v \equiv (v_i)$, let

$$f(x) = \sum_{j=1}^n v_j s_j(x).$$

We have $\langle f, f \rangle_w = \|v\|_2^2$ by Parseval's identity, and

$$\frac{v^* L P L^* v}{v^* v} = \frac{\int_{\mathbf{R}} |f(x)|^2 w(x) dx}{\int_{\mathbf{R}} |f(x)|^2 dx}.$$

Since $L P L^*$ and $K^{-1}P$ are similar, we have the claim. ■

Hence the distribution of $\lambda_i(K^{-1}P)$ depends on the capability of the functions $r_i(x)$ to approximate specific functions in $L^2(\mathbf{R})$. This topic seems not to have been dealt with in the literature, and it goes beyond the scope of the present paper. However, the following example shows that we cannot expect the values of $\lambda_i(K^{-1}P)$ to be fairly distributed in $[m, M]$, under general hypotheses.

In Figure 2 we plot the eigenvalues of preconditioned matrices $K_n^{-1}P_n$, of order n , where P_n is generated by $w(x) = 1 + (x^2 + 1)^{-1}$ and K_n is generated by the constant 1, under various configurations of the nodes. Remark that $1 < w(x) \leq 2$. In Figure 2a the nodes are the number 2^{i-1} , for $i = 1 \dots n$, and in Figure 2b their reciprocals. The nodes relative to Figure 2c are $2^{2^i - n - 1}$, and those in Figure 2d are the optimal nodes depicted in Figure 1d.

References

- [1] Ackner, R. and Kailath, T., On the Ptàk-Young generalization of the Schur-Cohn theorem. *IEEE Trans. Automat. Control*, 37 (1992), 1601–1604.
- [2] Ball, J.A.; Gohberg, I. and Rodman, L., *Interpolation of rational matrix functions*. Operator Theory: Advances and Applications, 45. Birkhäuser Verlag, Basel, 1990.
- [3] Borwein, J.M. and Lewis, A.S., Duality relationships for entropy-like minimization problems. *SIAM J. Control and Optimization*, 29 (1991), 325–338.

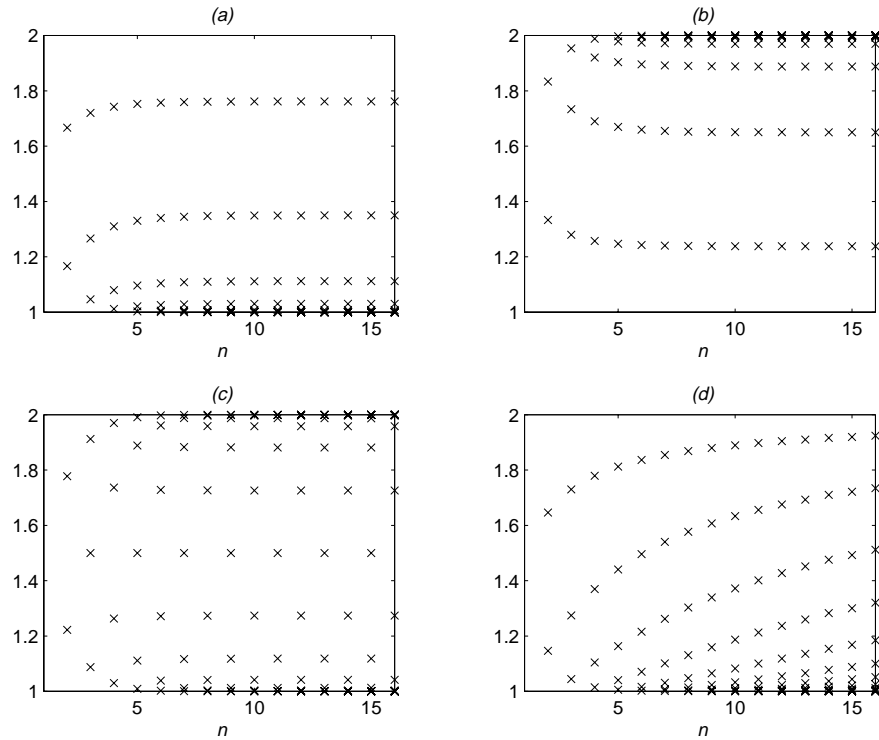


Figure 2: Eigenvalues of preconditioned Pick matrices

- [4] Fasino, D., Spectral properties of Hankel matrices and numerical solutions of finite moment problems. *J. Comput. Appl. Math.*, 65 (1995), 145–155.
- [5] Foias, C.; Frazho, A.E.; Gohberg, I. and Kaashoek, M.A., *Metric constrained interpolation, commutant lifting and systems*. Operator Theory: Advances and Applications, 100. Birkhäuser Verlag, Basel, 1998.
- [6] Frontini, M. and Tagliani, A., Maximum entropy in the finite Stieltjes and Hamburger moment problem. *J. Math. Physics*, 35 (1994), 6748–6756.
- [7] Gautschi, W. and Inglese, G., Lower bounds for the condition number of Vandermonde matrices. *Numer. Math.*, 52 (1988), 241–250.
- [8] Gohberg, I.; Kailath, T. and Olshevsky, V., Fast Gaussian elimination with partial pivoting for matrices with displacement structure. *Math. Comp.*, 64 (1995), no. 212, 1557–1576.
- [9] Kailath, T. and Sayed, A.H., Displacement structure: Theory and applications. *SIAM Rev.*, 37 (1985), 297–386.
- [10] Olshevsky, V. and Pan, V., *A unified superfast algorithm for boundary rational tangential interpolation problems and for inversion and factorization of dense structured matrices*. Proceedings of

the 39th Annual Symposium on Foundations of Computer Science, pp. 192–201, IEEE Computer Society Press, Los Alamitos, 1998.

- [11] Penzl, T., Eigenvalue decay bounds for solutions of Lyapunov equations: The symmetric case. *Systems and Control Letters* (submitted), 1999.
- [12] Sakhnovich, L.A., *Interpolation theory and its applications*. Mathematics and its Applications, 428. Kluwer Academic Publishers, Dordrecht, 1997.
- [13] Sarason, D., Nevanlinna-Pick interpolation with boundary data. *Integral Equations Operator Theory*, 30 (1998), 231–250.
- [14] Shohat, J.A. and Tamarkin, J.D., The problem of moments. *Amer. Math. Soc.*, New York 1943.
- [15] Taylor, J.M., The condition of Gram matrices and related problems. *Proc. of the Royal Soc. of Edinburgh*, 80a (1978), 45–56.
- [16] Tyrtyshnikov, E., How bad are Hankel matrices? *Numer. Math.*, 67 (1994), 261–269.