Submatrix Selection for Matrices with Orthonormal Columns

Tongren Xiao

Abstract

For a real matrix with three orthonormal columns, we obtain a lower bound for the largest squared smallest singular value among its three-row submatrices. For every $n \geq 6$, it improves the corresponding three-dimensional bound of Xu and the classical maximum-volume bound. The bound is asymptotic to $(5-\sqrt{13})/(2n)$. For six rows we obtain the larger bound given by the smallest zero of $70x^3 - 120x^2 + 57x - 6$. The proof combines a characteristic polynomial averaged with determinant weights, inequalities for squared minors, and incidence identities on three-element subsets. For arbitrary k, we also prove the bound $n^{-1/2}$ for an $n \times k$ matrix with orthonormal columns whose nonzero rows lie on at most $k+2$ distinct lines through the origin. We give an extremal construction for every $n > k$.

1 Introduction

For integers $n > k \geq 1$, let $[n] = \{1, \ldots, n\}$. If $U \in \mathbb{R}^{n \times k}$ and $I \subseteq [n]$, write U_I for the matrix formed by the rows indexed by I, in increasing order. Define

$$\Lambda_{n,k} = \min_{U^T U = I_k} \max_{\substack{I \subseteq [n] \\ |I| = k}} \sigma_{\min}(U_I)^2. \tag{1}$$

Here $\sigma_{\min}$ denotes the smallest singular value. The minimum exists because the set $\{U : U^T U = I_k\}$ is compact and the objective is continuous. Goreinov, Tyrtyshnikov, and Zamarashkin conjectured that $\Lambda_{n,k} \geq 1/n$ [1]. Their maximum-volume argument gives $\Lambda_{n,k} \geq [k(n-k)+1]^{-1}$. Nesterenko proved the case $(n,k) = (4,2)$ [2]; Sengupta and Pautov subsequently proved the real case $k = 2$ for every n [3].

Orthogonal complementation gives $\Lambda_{n,k} = \Lambda_{n,n-k}$. Indeed, complete U to an orthogonal matrix $[U\ V]$ and fix $|I| = k$. The equalities

$$U_I U_I^T = I_k - V_I V_I^T, \qquad V_{I^c}^T V_{I^c} = I_{n-k} - V_I^T V_I$$

show that both smallest eigenvalues equal $1 - \|V_I\|_2^2$. Consequently, $(5,3)$ is already a consequence of the two-column theorem.

The uniform interlacing method relates submatrix selection to zeros of averaged characteristic polynomials [7, Section 5]. Xu [5, Theorem 1.2] gives the three-dimensional interlacing bound explicitly. In the present notation, it is the smallest zero ρ_n of

$$f_n(x) = x^3 - \frac{9}{n}x^2 + \frac{18}{n(n-1)}x - \frac{6}{n(n-1)(n-2)}. \tag{2}$$

This is obtained by setting both the dimension and the selected size equal to three, and the number of available columns equal to n, in Xu's theorem, applied to U^T. Our estimate uses determinant weights instead of uniform weights. The stability-preserving operator theorem used in the selection step is due to Borcea and Brändén [8, Lemma 2.2].

Put $\Lambda_n = \Lambda_{n,3}$ and, for $n \geq 6$, define

$$\tau_n = \frac{5n^2 - 11n + 4 - \sqrt{(n^2 - 3n + 4)(13n^2 - 43n + 36)}}{2(n-2)(n^2 + n - 4)}. \tag{3}$$

Theorem 1.1. *For every $n \geq 6$,*

$$\Lambda_n \geq \tau_n > \frac{2}{3n-4} > \rho_n. \tag{4}$$

In particular,

$$\liminf_{n\to\infty} n\Lambda_n \geq \frac{5-\sqrt{13}}{2}.$$

For $n = 6$ one has $\Lambda_6 \geq \alpha_6 > \tau_6$, where α_6 is the unique zero in $(1/7, 3/20)$ of

$$g(x) = 70x^3 - 120x^2 + 57x - 6. \tag{5}$$

This zero is also the smallest real zero of g.

The comparisons with the classical bound and Xu's three-dimensional estimate are

$$\max\left\{\frac{1}{3n-8}, \rho_n\right\} < \frac{2}{3n-4} < \tau_n \qquad (n \geq 6). \tag{6}$$

Their proofs are included in Section 4. For $n = 6$, the uniform cubic gives $\rho_6 = (5-\sqrt{15})/10$, whereas our bound is α_6. Xu treats arbitrary dimensions and selection sizes; our comparison concerns selection of three rows from a matrix with three columns.

A line is a one-dimensional linear subspace.[1]

Theorem 1.2. *Let $n > k \geq 2$ and $U \in \mathbb{R}^{n\times k}$ satisfy $U^TU = I_k$. If the nonzero rows of U lie on at most $k+2$ distinct lines, then there is a set $I \subseteq [n]$, $|I| = k$, for which*

$$\sigma_{\min}(U_I) \geq n^{-1/2}. \tag{7}$$

For $k = 3$, Theorem 1.2 applies to matrices whose rows lie on at most five distinct lines. The proof uses the two-column theorem on the orthogonal complement of a matrix obtained by combining parallel rows. Proposition 5.2 gives equality examples for every $n > k$; for $k = 2$, equality is covered by Nesterenko's criterion [4]. The unrestricted case $(n, k) = (6, 3)$ is not settled here.

2 Minor Inequalities

For a real symmetric matrix H, the notation $H \succeq 0$ means that H is positive semidefinite, and $H \preceq K$ means $K - H \succeq 0$. Write $\|H\|_F^2 = \operatorname{tr}(H^TH)$ and let $\|H\|_2$ be its operator norm. Throughout Sections 2–4, $P = UU^T$ is a rank-three orthogonal projection on $\mathbb{R}^n$. Set

$$\begin{aligned}
&p_i = P_{ii}, \qquad b_{ij} = \det P_{\{i,j\},\{i,j\}}, \quad d_I = \det P_{I,I} \quad (|I| = 3),\\
&Q = \sum_i p_i^2, \quad B = \sum_{i<j} b_{ij}^2, \qquad C = \sum_{|I|=3} d_I^2.
\end{aligned} \tag{8}$$

All these numbers are nonnegative. Cauchy–Binet gives

$$\sum_I d_I = 1, \qquad \sum_{I\ni i} d_I = p_i, \qquad \sum_{I\supset\{i,j\}} d_I = b_{ij}. \tag{9}$$

To verify all three formulas, expand the same determinant in two ways:

$$\det\bigl(U^T(I_n + \operatorname{diag} z)U\bigr) = \sum_{|I|=3} d_I \prod_{i\in I}(1+z_i) = \sum_{J\subseteq[n]} \det P_{J,J} \prod_{j\in J} z_j.$$

The second equality uses $U^TU = I_3$, Sylvester's determinant identity, and the principal-minor expansion of $\det(I_n + \operatorname{diag}(z)P)$. The empty principal minor is defined to be one.

[1]The lines pass through the origin. Opposite vectors lie on the same line, and row lengths may differ. Zero rows do not contribute a direction.

Lemma 2.1. *Let $m \geq 3$. Suppose $q_{ij} = x_i y_j - x_j y_i$ for $x, y \in \mathbb{R}^m$, and extend the notation by $q_{ji} = -q_{ij}$ and $q_{ii} = 0$. Put $r_i = \sum_j q_{ij}^2$ and $\bar{r} = m^{-1} \sum_i r_i$. Then*

$$\sum_{i<j} q_{ij}^4 \geq \frac{3}{4m} \sum_i r_i^2, \tag{10}$$

and

$$\sum_{i<j} q_{ij}^4 - \frac{3}{4m} \sum_i r_i^2 \geq \frac{1}{4m} \sum_i (r_i - \bar{r})^2. \tag{11}$$

Proof. The zero case is immediate. Gram–Schmidt and homogeneity reduce the proof to orthonormal x, y. Define $H = xx^T + yy^T$ and let D be the symmetric matrix with $D_{ij} = q_{ij}^2$ for $i \neq j$ and zero diagonal. Then $r_i = H_{ii}$, $D\mathbf{1} = r$, and

$$D = rr^T - H \circ H,$$

where $\circ$ denotes entrywise multiplication. Let W have rows $(x_i^2, \sqrt{2} x_i y_i, y_i^2)$. Then $H \circ H = WW^T$ and $r = W(1, 0, 1)^T$. Thus $\operatorname{rank} D \leq 3$, and D has at most one positive eigenvalue. Since $D \neq 0$ and $\operatorname{tr} D = 0$, its spectrum, with zeros appended, is $\lambda, -a, -b$, where $\lambda > 0$, $a, b \geq 0$ and $a + b = \lambda$. Let $\theta = \max(a, b)/\lambda \in [1/2, 1]$. Choose an orthonormal eigenbasis v_j of D/λ, with eigenvalues $\xi_j \in [-\theta, 1]$, and set

$$w_j = \frac{|v_j^T \mathbf{1}|^2}{m}, \qquad m_1 = \sum_j w_j \xi_j, \qquad m_2 = \sum_j w_j \xi_j^2.$$

Then $\sum_j w_j = 1$, $\|r\|^2 = m\lambda^2 m_2$ and $\|r - \bar{r}\mathbf{1}\|^2 = m\lambda^2(m_2 - m_1^2)$. The scalar inequality $s^2 \leq (1 - \theta)s + \theta$ on $[-\theta, 1]$ implies

$$4m_2 - m_1^2 \leq 4(1 - \theta)m_1 + 4\theta - m_1^2 \leq 4(\theta^2 - \theta + 1).$$

Since

$$\sum_{i<j} q_{ij}^4 = \tfrac{1}{2} \|D\|_F^2 = \lambda^2(\theta^2 - \theta + 1),$$

this proves (11). Inequality (10) follows because the right-hand side of (11) is nonnegative. □

Lemma 2.2. *For $n \geq 4$, the quantities in (8) satisfy*

$$B \leq 2(n-1)C, \qquad B \leq \frac{3(n-1)}{2} C + \frac{Q}{2(n-1)}. \tag{12}$$

Proof. Fix i with $p_i > 0$. By a right orthogonal change of coordinates, make the ith row of U equal to $(\sqrt{p_i}, 0, 0)$. For $j, \ell \neq i$, the signed maximal minor with rows i, j, ℓ is

$$q_{j\ell} = \sqrt{p_i}(U_{j2}U_{\ell 3} - U_{j3}U_{\ell 2}).$$

These are the coefficients of a decomposable two-form on $\mathbb{R}^{n-1}$, meaning exactly the representation in Lemma 2.1. Their squares are $d_{\{i,j,\ell\}}$. The row sums of the matrix of squared coefficients are b_{ij} by (9), and their sum is $2p_i$. Put $C_i = \sum_{I \ni i} d_I^2$ and $B_i = \sum_{j \neq i} b_{ij}^2$. With $m = n - 1$, Lemma 2.1 gives

$$C_i \geq \frac{3}{4m} B_i, \qquad C_i \geq \frac{1}{m} B_i - \frac{p_i^2}{m^2}.$$

If $p_i = 0$, the ith row of U vanishes and both inequalities still hold. Summing and using $\sum_i C_i = 3C$, $\sum_i B_i = 2B$ gives (12). □

Lemma 2.3. *For $n \geq 6$, the quantities in (8) satisfy*

$$\frac{4}{n-2}Q - \frac{18}{(n-1)(n-2)} \leq B \leq (n-4)C + \frac{2}{n-3}Q - \frac{6}{(n-2)(n-3)}. \tag{13}$$

Proof. Let A be the vertex–triple incidence matrix and M the pair–triple incidence matrix: $A_{i,I} = 1$ if $i \in I$ and zero otherwise, and $M_{e,I} = 1$ if $e \subset I$ and zero otherwise. Let E be the vertex–pair incidence matrix. A matrix denoted by J below has all entries equal to one; its size is specified by each identity. Counting triples and pairs gives

$$AA^T = \binom{n-2}{2} I_n + (n-2)J_n, \qquad EE^T = (n-2)I_n + J_n,$$

$$MM^T = (n-4)I_{\binom{n}{2}} + E^T E, \qquad EM = 2A, \qquad AM^T = (n-3)E + J_{n,\binom{n}{2}}.$$

In $\mathbb{R}^{\binom{n}{3}}$ define

$$V_0 = \mathbb{R}\mathbf{1}, \quad V_1 = \operatorname{im} A^T \cap V_0^\perp, \quad V_2 = \operatorname{im} M^T \cap (V_0 \oplus V_1)^\perp, \quad V_3 = \ker M.$$

The first three identities show that A and M have full row rank; $EM = 2A$ shows $\operatorname{im} A^T \subseteq \operatorname{im} M^T$. These facts give the orthogonal decomposition into V_0, V_1, V_2, V_3. Decompose $\mathbb{R}^{\binom{n}{2}}$ into constants, $E^T\mathbf{1}^\perp$, and $\ker E$. The identity for MM^T gives respective eigenvalues $3(n-2), 2(n-3), n-4$. The last two identities identify their images under M^T with V_0, V_1, V_2, respectively. Therefore

	V_0	V_1	V_2	V_3
$A^T A$	$3\binom{n-1}{2}$	$\binom{n-2}{2}$	0	0
$M^T M$	$3(n-2)$	$2(n-3)$	$n-4$	0
$J_{\binom{n}{3}}$	$\binom{n}{3}$	0	0	0

For $d = (d_I)$, formulas (9) give $\|Ad\|^2 = Q$, $\|Md\|^2 = B$, $\|d\|^2 = C$ and $d^T J d = 1$. Evaluating the table yields

$$B - \frac{4}{n-2}Q + \frac{18}{(n-1)(n-2)} = (n-4)\|\Pi_{V_2} d\|^2,$$

$$(n-4)C + \frac{2}{n-3}Q - \frac{6}{(n-2)(n-3)} - B = (n-4)\|\Pi_{V_3} d\|^2,$$

where Π_V is orthogonal projection onto V. Both right-hand sides are nonnegative. □

3 Averaged Characteristic Polynomials

A real polynomial is stable if it is nonzero whenever all its variables have positive imaginary parts. It is multiaffine if its degree in each variable is at most one. A univariate real stable polynomial has only real zeros. We use the following form of the Borcea–Brändén symbol theorem [8, Lemma 2.2]: a linear operator T on multiaffine polynomials preserves stability, allowing the zero polynomial, if $T[\prod_i (z_i + w_i)]$ is stable in the output variables and the w_i. We also use preservation under differentiation and real specialization [8, Lemma 1.7].

Lemma 3.1. *Let μ be a probability measure on subsets of $[N]$ such that $g(z) = \sum_S \mu(S) \prod_{i \in S} z_i$ is stable. Let H_0 be a real symmetric $r \times r$ matrix and let $H_i = a_i a_i^T$ for $a_i \in \mathbb{R}^r$. Then*

$$\sum_S \mu(S) \det\left(xI_r - H_0 - \sum_{i \in S} H_i\right)$$

has only real zeros.

Proof. The polynomial $F(x,y) = \det(xI_r - H_0 + \sum_i y_i H_i)$ is multiaffine in y because the H_i have rank at most one. It is stable: the imaginary part of the matrix is positive definite when all variables are in the upper half-plane. Define

$$Th = F(x, -\partial_{z_1}, \dots, -\partial_{z_N}) h(z)\big|_{z=\mathbf{1}}.$$

Its symbol is

$$T\left[\prod_i (z_i + w_i)\right] = \prod_i (1+w_i) F\left(x, -\frac{1}{1+w_1}, \dots, -\frac{1}{1+w_N}\right).$$

Each $-(1+w_i)^{-1}$ has positive imaginary part when w_i does. The symbol is therefore stable, so Tg is stable or zero. Writing $y^J = \prod_{j\in J} y_j$ and $F(x,y) = \sum_J F_J(x) y^J$, we have

$$Tg = \sum_J (-1)^{|J|} F_J(x) \sum_{S \supseteq J} \mu(S) = \sum_S \mu(S) F(x, -\mathbf{1}_S).$$

Here $\mathbf{1}_S$ is the indicator vector of S. The resulting polynomial is monic and hence not zero. □

Lemma 3.2. *The polynomial*

$$\chi(x) = \sum_{|I|=3} d_I \det(xI_3 - U_I^T U_I) = x^3 - Qx^2 + Bx - C \tag{14}$$

has zeros $0 < r_1 \le r_2 \le r_3 \le 1$. *Some triple* I *satisfies* $\sigma_{\min}(U_I)^2 \ge r_1$.

Proof. The weights d_I sum to one, and

$$\sum_I d_I z^I = \det\bigl(U^T \operatorname{diag}(z) U\bigr)$$

is stable because its matrix has positive definite imaginary part. Lemma 3.1 gives real-rootedness of χ. The trace of $U_I^T U_I$ is $\sum_{i\in I} p_i$; its second elementary symmetric function is $\sum_{\{i,j\}\subset I} b_{ij}$ by Cauchy–Binet. Thus (9) gives the coefficient formula in (14). For $x < 0$ each determinant in the average is negative, and for $x > 1$ each is positive, because $0 \preceq U_I^T U_I \preceq I_3$. There are no zeros outside $[0,1]$. Also $C > 0$, so zero is not a root.

For selection, condition successively on inclusion or exclusion of coordinates in the random triple. These operations replace its generating polynomial by a derivative or a specialization, followed by normalization; nonzero conditional polynomials remain stable. At a node with both children of positive probability, any convex combination of the child distributions can be obtained by multiplying the weights of subsets containing that coordinate by a nonnegative scalar, and taking a limit for the endpoint that conditions on inclusion. This is a nonnegative rescaling of a variable in the conditional generating polynomial. Lemma 3.1, with the already selected vectors included in H_0, shows that every convex combination of the two child characteristic polynomials is real-rooted.

The common-interlacing criterion [6, Lemma 3.5] now applies. Here common interlacing means that there are two real numbers $\eta_1 \le \eta_2$ such that the zeros $\beta_1 \le \beta_2 \le \beta_3$ of each child satisfy $\beta_1 \le \eta_1 \le \beta_2 \le \eta_2 \le \beta_3$. For monic polynomials with common interlacing, the smallest zero of a convex combination lies between the smallest zeros of its children. Choose a child whose smallest zero is at least that of its parent. Nodes with only one possible child require no choice. After all coordinates have been conditioned on, the polynomial is $\det(xI_3 - U_I^T U_I)$ for a triple I, proving the assertion. □

4 Proof of the Three-Column Bounds

Proof of the bound $\Lambda_n \geq \tau_n$ in Theorem 1.1. Put $d_n = n^2 + n - 4$ and

$$h_n(t) = (n-2)d_n t^2 - (5n^2 - 11n + 4)t + 3n - 4.$$

Its discriminant is $\Delta_n = (n^2 - 3n + 4)(13n^2 - 43n + 36)$, and τ_n is its smaller zero. Define

$$\mathcal{U} = B - (n-4)C - \frac{2}{n-3}Q + \frac{6}{(n-2)(n-3)}, \qquad \mathcal{S} = B - \frac{3(n-1)}{2}C - \frac{Q}{2(n-1)},$$

and $\mathcal{R} = C + Q - B - 1$. Lemmas 2.2 and 2.3 give $\mathcal{U}, \mathcal{S} \leq 0$. Lemma 3.2 gives $-\mathcal{R} = \chi(1) = \prod_j (1 - r_j) \geq 0$. The following identity is obtained by collecting coefficients of Q, B, C:

$$\begin{aligned} \chi(t) &= \lambda_U(t)\mathcal{U} + \lambda_S(t)\mathcal{S} + \lambda_R(t)\mathcal{R} \\ &\quad + \frac{(t-1)h_n(t)}{(n-2)d_n}, \end{aligned} \tag{15}$$

where

$$\begin{aligned} \lambda_U(t) &= -\frac{(n-3)(t-1)((3n^2 - 8n + 5)t - 2n + 3)}{(n-5)d_n}, \\ \lambda_S(t) &= \frac{2(n-1)(t-1)((n-3)t - 1)}{d_n}, \\ \lambda_R(t) &= -\frac{N_n(t)}{(n-5)d_n}, \\ N_n(t) &= (n^3 + n^2 - 17n + 15)t^2 + (-5n^2 + 5n + 6)t + 3n - 1. \end{aligned} \tag{16}$$

We verify all multiplier signs. Set $\ell_n = (2n-3)/(3n^2 - 8n + 5)$. Substitution gives

$$h_n(\ell_n) = \frac{(n-2)^3 d_n}{(n-1)^2(3n-5)^2} > 0, \qquad h_n(1/(n-3)) = -\frac{(n-4)d_n}{(n-3)^2} < 0.$$

Since $0 < \ell_n < 1/(n-3)$, we have $\ell_n < \tau_n < 1/(n-3) < 1$, and hence $\lambda_U(\tau_n), \lambda_S(\tau_n) > 0$.

For the remaining sign, put

$$a = \frac{5 - \sqrt{13}}{2}, \qquad L_n = \frac{5a}{5n-6}, \qquad u = n - 6 \geq 0.$$

Using $18/5 < \sqrt{13} < 181/50$, direct expansion gives

$$\begin{aligned} (5n-6)^2 h_n(L_n) &= (545 - 150\sqrt{13})u^2 + (4388 - 1210\sqrt{13})u \\ &\quad + 8764 - 2420\sqrt{13} > 2u^2 + \frac{39}{5}u + \frac{18}{5} > 0, \\ 2(5n-6)^2 N_n(L_n) &= (1440 - 400\sqrt{13})u^2 + (9541 - 2675\sqrt{13})u \\ &\quad + 11559 - 3345\sqrt{13} < -89u - 483 < 0. \end{aligned} \tag{17}$$

Since $0 < a < 3/4$, we have $L_n < 15/[4(5n-6)] < 1/(n-1)$. Together with

$$h_n(1/(n-1)) = -\frac{(n-2)(n^2 - 3n + 4)}{(n-1)^2} < 0,$$

we obtain $L_n < \tau_n < 1/(n-1)$. The quadratic N_n is convex because its leading coefficient is $(n-3)(n-1)(n+5) > 0$. Its other endpoint value is

$$N_n(1/(n-1)) = -\frac{n^2 - 3n + 8}{n-1} < 0.$$

Convexity therefore gives $N_n(\tau_n) < 0$ and $\lambda_R(\tau_n) > 0$. Equation (15) now implies $\chi(\tau_n) \le 0$. By Vieta's formulas and (12),

$$\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} = \frac{B}{C} \le 2(n-1). \tag{18}$$

If $r_1 < \tau_n$ and $r_2 \le \tau_n$, the first two terms already exceed $2/\tau_n > 2(n-1)$, a contradiction. Thus $r_1 < \tau_n$ would imply $r_1 < \tau_n < r_2 \le r_3$ and hence $\chi(\tau_n) > 0$, again a contradiction. Therefore $r_1 \ge \tau_n$, and Lemma 3.2 proves $\Lambda_n \ge \tau_n$. □

Proof of the comparisons and limit in Theorem 1.1. Let $v_n = 2/(3n-4)$. One has

$$h_n(v_n) = \frac{n(n-4)(n-2)}{(3n-4)^2} > 0, \qquad v_n < 1/(n-3),$$

and $h_n(1/(n-3)) < 0$, so $v_n < \tau_n$. For the classical bound,

$$v_n - \frac{1}{3n-8} = \frac{3(n-4)}{(3n-4)(3n-8)} > 0.$$

For the uniform cubic (2),

$$f_n(0) < 0, \qquad f_n(v_n) = \frac{2(n-4)(31n^2-86n+60)}{n(n-1)(n-2)(3n-4)^3} > 0.$$

It has a zero in $(0, v_n)$, so $\rho_n < v_n$. This proves (6). Finally, multiplying (3) by n and dividing numerator and denominator by their leading powers of n gives $n\tau_n \to (5-\sqrt{13})/2$. This proves the stated limit bound. □

Proof of the six-row bound in Theorem 1.1. The derivative $g'(x) = 210x^2 - 240x + 57$ is positive on $[0, 3/20]$: it decreases there and $g'(3/20) > 0$. Since $g(1/7) = -5/49$ and $g(3/20) = 69/800$, there is a unique zero $a = \alpha_6$ in $(1/7, 3/20)$ and no smaller nonnegative zero. All terms of $g(x)$ are negative for $x < 0$, so a is its smallest zero.

By (12) and (13),

$$Q - \frac{9}{10} \le B, \qquad B \le 2C + \frac{2}{3}Q - \frac{1}{2}, \qquad B \le \frac{15}{2}C + \frac{Q}{10}. \tag{19}$$

Set

$$\mathcal{L} = Q - \frac{9}{10} - B, \quad \mathcal{U} = B - 2C - \frac{2}{3}Q + \frac{1}{2}, \quad \mathcal{S} = B - \frac{15}{2}C - \frac{Q}{10}.$$

All three are nonpositive by (19). Coefficient comparison gives

$$\begin{aligned} \chi(t) = {} & \frac{-165t^2 + 144t - 17}{21}\mathcal{L} + \frac{-75t^2 + 75t - 9}{7}\mathcal{U} \\ & + \frac{60t^2 - 60t + 10}{21}\mathcal{S} + \frac{g(t)}{70}. \end{aligned} \tag{20}$$

Each coefficient of $\mathcal{L}, \mathcal{U}, \mathcal{S}$ is positive on $[1/7, 3/20]$. The first two numerators increase on this interval and their values at $1/7$ are $10/49$ and $9/49$, respectively. The third numerator decreases and its value at $3/20$ is $47/20$. Evaluating (20) at $t = a$ gives $\chi(a) \le 0$. If $r_1 < a$ and $r_2 \le a$, then $1/r_1 + 1/r_2 > 2/a > 10$, contradicting (18). Thus $r_1 < a$ would force $\chi(a) > 0$. We obtain $r_1 \ge a$, and Lemma 3.2 proves the result. Finally, $h_6(3/20) = -7/25 < 0$ and the preceding bounds give $1/7 < \tau_6 < 3/20$. Reducing g modulo h_6 yields

$$g(\tau_6) = \frac{137 - 1209\tau_6}{2888} < 0.$$

Since g is increasing on $[0, 3/20]$, we obtain $\tau_6 < a$. □

5 Matrices with at Most $k+2$ Row Directions

Lemma 5.1. *Let $r \in \{1,2\}$, let $y_1,\dots,y_m \in \mathbb{R}^r$, and let $p_i > 0$ with $\sum_i p_i \le 1$. If $H = \sum_i p_i y_i y_i^T \succ 0$, there is $J \subseteq [m]$, $|J| = r$, such that $\sum_{j\in J} y_j y_j^T \succeq H$.*

Proof. For $r=1$, select an index maximizing y_i^2. For $r=2$, congruence by $H^{-1/2}$ reduces to $H = I_2$. First suppose $p_i = N_i/N$ for positive integers N_i and N. Increase the common denominator if necessary so that $N > 2$. Form an $N \times 2$ matrix with N_i copies of $y_i^T/\sqrt{N}$ and $N - \sum_i N_i$ zero rows. Its columns are orthonormal. The theorem of Sengupta and Pautov [3] selects two rows with squared smallest singular value at least $1/N$. Neither is zero, and they cannot come from the same index because the selected matrix is nonsingular. Their indices j,ℓ satisfy $y_j y_j^T + y_\ell y_\ell^T \succeq I_2$.

For real weights, choose positive rational $p_i^{(s)} \to p_i$ with $\sum_i p_i^{(s)} \le 1$. Then $H_s = \sum_i p_i^{(s)} y_i y_i^T \to I_2$ and $H_s \succ 0$ for large s. Apply the rational case to $H_s^{-1/2} y_i$. A fixed pair is selected along a subsequence, because there are only finitely many pairs. Congruence back gives $y_j y_j^T + y_\ell y_\ell^T \succeq H_s$; passage to the limit proves the assertion. □

Proof of Theorem 1.2. Partition the nonzero rows into m distinct lines, where $k \le m \le k+2$. In group j write the rows as $a_{j\ell} v_j^T$, with $\|v_j\| = 1$ and $a_{j\ell} \ne 0$, $1 \le \ell \le m_j$. Set

$$s_j = \sum_\ell a_{j\ell}^2, \qquad w_j = \sqrt{s_j}\, v_j, \qquad c_j = \frac{\max_\ell a_{j\ell}^2}{s_j}.$$

The matrix W with rows w_j^T satisfies $W^T W = I_k$, and

$$c_j \ge \frac{1}{m_j}, \qquad \sum_j \frac{1}{c_j} \le \sum_j m_j \le n. \tag{21}$$

If $m = k$, then W is orthogonal. Choosing a longest row from every group gives a Gram matrix $\operatorname{diag}(c_j) \succeq I_k/n$.

Suppose $m = k + r$ with $r \in \{1,2\}$. Complete W to an orthogonal matrix $[W\ Z]$, where $Z \in \mathbb{R}^{m\times r}$ has rows z_j^T. Then $WW^T = I_m - ZZ^T$ and $\sum_j z_j z_j^T = I_r$. Put $t = 1/n$. Since each group omits at least one other nonempty group, $m_j < n$ and $c_j > t$. Define

$$b_j = \frac{t}{c_j - t}, \qquad y_j = \sqrt{1+b_j}\, z_j, \qquad p_j = \frac{t}{c_j}.$$

The weights satisfy $\sum_j p_j \le 1$ by (21), and

$$H = \sum_j p_j y_j y_j^T = \sum_j b_j z_j z_j^T \succ 0.$$

Positive definiteness holds because all $b_j > 0$ and Z has full column rank. Lemma 5.1 gives $|J| = r$ such that $\sum_{j\in J}(1+b_j) z_j z_j^T \succeq H$. Let $I = [m] \setminus J$ and select a longest row from every group indexed by I. Up to row signs, its Gram matrix is $D(I_k - Z_I Z_I^T)D$, where $D = \operatorname{diag}(\sqrt{c_i} : i \in I)$. The Schur-complement criterion gives

$$\begin{aligned} D(I_k - Z_I Z_I^T)D \succeq tI_k &\iff Z_I^T D(D^2 - tI_k)^{-1} D Z_I \preceq I_r \\ &\iff \sum_{i\in I}(1+b_i) z_i z_i^T \preceq I_r. \end{aligned} \tag{22}$$

For justification, $A - XX^T \succeq 0$ with $A \succ 0$ is equivalent to $X^T A^{-1} X \preceq I$. This follows by congruence with $A^{-1/2}$ and equality of the nonzero eigenvalues of YY^T and $Y^T Y$, where $Y = A^{-1/2}X$. The last inequality of (22) follows from

$$\sum_{i\in I}(1+b_i) z_i z_i^T = I_r + H - \sum_{j\in J}(1+b_j) z_j z_j^T \preceq I_r.$$

This proves (7). $\square$

Proposition 5.2. *For every $n > k \geq 1$, there is $U \in \mathbb{R}^{n\times k}$ with $U^T U = I_k$ such that*

$$\max_{|I|=k} \sigma_{\min}(U_I)^2 = \frac{1}{n}. \tag{23}$$

For $k \geq 2$, its nonzero rows lie on $k+1$ distinct lines.

Proof. For $k = 1$, take all rows equal to $n^{-1/2}$. Let $k \geq 2$. Choose $u_1, \ldots, u_k \in \mathbb{R}^{k-1}$ with

$$\sum_j u_j u_j^T = I_{k-1}, \quad \sum_j u_j = 0, \quad \|u_j\|^2 = \frac{k-1}{k}, \quad u_i^T u_j = -\frac{1}{k}\ (i \neq j).$$

Such vectors are the rows of a matrix whose columns form an orthonormal basis of $\mathbf{1}^\perp \subset \mathbb{R}^k$. Set $a = \sqrt{(n-1)/(n(n-k))}$ and $b = 1/\sqrt{kn}$. Take $n-k$ rows equal to $x = (a, 0, \ldots, 0)$, and the remaining rows $y_j = (b, u_j)$, $1 \leq j \leq k$. Their outer products sum to I_k.

The submatrix consisting of all the y_j has column Gram matrix $\operatorname{diag}(1/n, 1, \ldots, 1)$. Any submatrix with two copies of x is singular. For a submatrix with one x and all y_j except y_q, the column Gram matrix equals the identity on the $(k-2)$-dimensional space orthogonal to the first coordinate and u_q. On the remaining two-dimensional space it is

$$H_q = \begin{pmatrix} a^2 + (k-1)b^2 & -b\sqrt{(k-1)/k} \\ -b\sqrt{(k-1)/k} & 1/k \end{pmatrix}.$$

The diagonal entries of $H_q - I_2/n$ are positive, and

$$\left(a^2 + (k-1)b^2 - \frac{1}{n}\right)\left(\frac{1}{k} - \frac{1}{n}\right) = \frac{k-1}{k^2 n} = b^2 \frac{k-1}{k}.$$

Thus $H_q - I_2/n$ is positive semidefinite with determinant zero, so the smallest eigenvalue of H_q is exactly $1/n$. This exhausts all submatrices and proves (23). The u_j are distinct, all y_j have the same positive first coordinate, and no u_j is zero. Hence $x, y_1, \ldots, y_k$ determine $k+1$ distinct lines. $\square$

References

[1] S. A. Goreinov, E. E. Tyrtyshnikov, and N. L. Zamarashkin, A theory of pseudoskeleton approximations, *Linear Algebra Appl.* **261** (1997), 1–21. `https://doi.org/10.1016/S0024-3795(96)00301-1`.

[2] Y. Nesterenko, Submatrices with the best-bounded inverses: revisiting the hypothesis, arXiv:2303.07492 (2023). `https://arxiv.org/abs/2303.07492`.

[3] R. Sengupta and M. Pautov, On the submatrices with the best-bounded inverses, arXiv:2604.05944 (2026). `https://arxiv.org/abs/2604.05944`.

[4] Y. Nesterenko, Submatrices with the best-bounded inverses: the equality criteria for $\mathbb{R}^{n\times 2}$, arXiv:2604.14050 (2026). `https://arxiv.org/abs/2604.14050`.

[5] Z. Xu, New lower bounds for the minimum singular value in matrix selection, arXiv:2508.10452 (2025). `https://arxiv.org/abs/2508.10452`.

[6] A. W. Marcus, D. A. Spielman, and N. Srivastava, Interlacing families II: mixed characteristic polynomials and the Kadison–Singer problem, *Ann. of Math.* **182** (2015), 327–350. `https://doi.org/10.4007/annals.2015.182.1.8`.

[7] A. W. Marcus, D. A. Spielman, and N. Srivastava, Interlacing families III: sharper restricted invertibility estimates, *Israel J. Math.* **247** (2022), 519–546. `https://doi.org/10.1007/s11856-021-2277-z`.

[8] J. Borcea and P. Brändén, The Lee–Yang and Pólya–Schur programs. I. Linear operators preserving stability, *Invent. Math.* **177** (2009), 541–569. `https://arxiv.org/abs/0809.0401`.