Local Maxima of the Entrywise ℓ_4 Norm on the Orthogonal Group

Dian Jin
Department of Statistics and Data Science
National University of Singapore
dianjin@nus.edu.sg

Abstract

We classify the local maximizers of the entrywise fourth-power objective

$$Q \longmapsto \|Q\|_4^4 = \sum_{i,j=1}^{r} q_{ij}^4$$

over the real orthogonal group $\mathcal{O}(r)$. We prove that the signed permutation matrices are the only local maximizers, and hence the only global maximizers, in every dimension. More strongly, every other stationary point has an explicit rank-two tangent direction with strictly positive second variation. The proof is based on a maximum-entry pivot for the orthostochastic matrix $Q^{\circ 2}$: the associated full Riemannian Hessian can be evaluated exactly and is positive at a largest nonunit squared entry. The argument is self-contained and handles zeros, repeated magnitudes, reducible support, and Hadamard-type stationary points.

Keywords. Orthogonal group; entrywise ℓ_4 norm; Riemannian Hessian; local maxima; Hadamard matrices.

Throughout, powers marked by "$\circ$" are entrywise. In particular, $Q^{\circ 2} = (q_{ij}^2)$ and $Q^{\circ 3} = (q_{ij}^3)$. The norm in the objective is also entrywise:

$$f_r(Q) = \|Q\|_4^4 = \sum_{i,j=1}^{r} q_{ij}^4, \qquad Q \in \mathcal{O}(r).$$

1 Problem formulation and main result

The orthogonal group is compact but nonconvex, and entrywise norm objectives have a markedly different geometry from spectral objectives. The squared Frobenius norm is constant on $\mathcal{O}(r)$, whereas the fourth-power objective

$$f_r(Q) = \|Q\|_4^4, \qquad Q \in \mathcal{O}(r), \tag{1.1}$$

rewards concentration of mass in individual entries. This paper determines its complete local-maximality landscape over the real square orthogonal group.

The signed permutation matrices are exactly the local maximizers, and the maximum value is r. More strongly, every stationary point that is not a signed permutation has an explicit feasible rank-two direction with strictly positive second variation. Thus it has a strict ascent direction for the maximization problem.

Dimension	Local maximizers	Maximum	Other stationary points
$r = 1$	$[1]$ and $[-1]$	1	none
$r = 2$	signed permutations only	2	strict ascent direction exists
$r \geq 3$	signed permutations only	r	strict ascent direction exists

Here "strict ascent direction" means that the largest eigenvalue of the constrained Hessian is positive. This is sometimes called the strict-saddle property with the maximization sign convention. It does not assert that the Hessian is necessarily indefinite: a local minimum also has positive curvature directions.

2 Related work

The Riemannian geometry of optimization with orthogonality constraints is classical; standard treatments include the geometric analysis of the Stiefel and Grassmann manifolds by Edelman, Arias, and Smith [1] and the monograph of Absil, Mahony, and Sepulchre [2]. We use this framework only to derive the exact stationarity condition and the second variation along $Qe^{t\Omega}$.

Our problem also belongs to the study of entrywise ℓ_p norm landscapes on the orthogonal group. In particular, the almost-Hadamard program investigates local extrema of such norms, especially in connection with the ℓ_1 norm, Hadamard matrices, and more general exponents; see Banica and Nechita [3]. The $p = 4$ maximization problem has the opposite global extremizers from the normalized Hadamard configurations: signed permutations maximize $\|Q\|_4^4$. The result here gives a direct local classification for this fourth-power objective, without relying on a classification of stationary orthogonal matrices or on a relaxation to the Birkhoff polytope.

3 Main theorem

Theorem 1 (Complete local-maximum classification)**.** *For every integer $r \geq 1$, let*

$$f_r(Q) = \sum_{i,j=1}^{r} q_{ij}^4, \qquad Q \in \mathcal{O}(r).$$

Then:

(i) *$f_r(Q) \leq r$, with equality exactly at the signed permutation matrices;*

(ii) *every signed permutation is a strict local maximum (with the usual vacuous isolated-point interpretation when $r = 1$);*

(iii) *if Q is stationary and is not a signed permutation, then there is an explicit nonzero skew-symmetric matrix Ω such that*

$$\frac{d^2}{dt^2} f_r\left(Qe^{t\Omega}\right)\bigg|_{t=0} > 0.$$

Consequently, the signed permutation matrices are exactly the local maximizers of f_r on $\mathcal{O}(r)$.

The escaping matrix Ω is constructed in Section 6 from a largest entry of the squared-entry matrix $Q^{\circ 2}$.

4 First-order condition

The tangent space at $Q \in \mathcal{O}(r)$ is

$$T_Q\mathcal{O}(r) = \{Q\Omega : \Omega^\top = -\Omega\}.$$

Fix a skew-symmetric Ω and use the feasible curve

$$Q(t) = Qe^{t\Omega}.$$

Since $Q'(0) = Q\Omega$ and the Euclidean gradient is $4Q^{\circ 3}$,

$$\begin{aligned}
\frac{d}{dt} f_r(Q(t))\bigg|_{t=0} &= 4\left\langle Q^{\circ 3}, Q\Omega\right\rangle_F \\
&= 4\operatorname{tr}\Big((Q^{\circ 3})^\top Q\Omega\Big) \\
&= 4\operatorname{tr}(A^\top\Omega), \qquad A := Q^\top Q^{\circ 3}.
\end{aligned}$$

The orthogonal complement of the skew-symmetric matrices under the Frobenius inner product is the space of symmetric matrices. Therefore

$$\boxed{Q \text{ is stationary} \quad\Longleftrightarrow\quad A = Q^\top Q^{\circ 3} \text{ is symmetric.}}$$

This uses every tangent direction, not merely coordinate rotations.

For comparison, rotate columns a, b by

$$q_a(t) = q_a\cos t + q_b\sin t, \qquad q_b(t) = -q_a\sin t + q_b\cos t.$$

Direct differentiation of the affected two columns gives

$$f'_{ab}(0) = 4\sum_i q_{ia}q_{ib}(q_{ia}^2 - q_{ib}^2), \tag{1}$$

$$f''_{ab}(0) = 4\sum_i \left(6q_{ia}^2q_{ib}^2 - q_{ia}^4 - q_{ib}^4\right). \tag{2}$$

Equation (1) vanishes for every pair at a stationary point, consistently with the full matrix condition above. Formula (2) is only a sanity check; the proof below does not assume that coordinate Givens directions suffice.

5 Second-order formula

Because $Q''(0) = Q\Omega^2$, direct differentiation gives

$$f''_r(0) = 12\sum_{k,l} q_{kl}^2(Q\Omega)_{kl}^2 + 4\left\langle Q^{\circ 3}, Q\Omega^2\right\rangle_F.$$

At a stationary point $A = A^\top$, and hence

$$\left\langle Q^{\circ 3}, Q\Omega^2\right\rangle_F = \operatorname{tr}(A\Omega^2) = -\operatorname{tr}(A\Omega^\top\Omega).$$

Thus

$$\boxed{f''_r(0) = 4\mathcal{H}_Q(\Omega), \qquad \mathcal{H}_Q(\Omega) = 3\sum_{k,l} q_{kl}^2(Q\Omega)_{kl}^2 - \operatorname{tr}(A\Omega^\top\Omega).} \tag{3}$$

For maximization, a second-order necessary condition at a local maximum is $\mathcal{H}_Q(\Omega) \le 0$ for every skew Ω. Hence any Ω with $\mathcal{H}_Q(\Omega) > 0$ is a strict feasible escape certificate.

6 Main proof: the maximum-entry pivot

Let Q be stationary. Define

$$M = Q^{\circ 2}, \qquad s_i = \sum_l M_{il}^2, \qquad t_j = \sum_k M_{kj}^2, \qquad C = MM^\top M.$$

Orthogonality of the rows and columns of Q shows that M is doubly stochastic:

$$M_{ij} \geq 0, \qquad \sum_j M_{ij} = 1, \qquad \sum_i M_{ij} = 1.$$

6.1 An exact pivot identity

Fix indices (i, j) for which

$$p = M_{ij} = q_{ij}^2 < 1, \qquad q = q_{ij}.$$

Regard the transpose of row i as the unit column vector

$$x = Q^\top e_i.$$

Then $Qx = e_i$ and $x_j = q$. Set

$$v = \frac{e_j - qx}{\sqrt{1-p}}, \qquad \Omega = xv^\top - vx^\top. \tag{4}$$

Indeed,

$$x^\top v = 0, \qquad \|e_j - qx\|_2^2 = 1 - p,$$

so x, v are orthonormal. Thus Ω is nonzero, skew-symmetric, rank two, and generates the feasible curve $Qe^{t\Omega}$.

Since

$$Qv = \frac{Qe_j - qe_i}{\sqrt{1-p}},$$

a direct cancellation in $Q\Omega = e_i v^\top - (Qv)x^\top$ gives

$$(Q\Omega)_{kl} = \frac{\delta_{ki}\delta_{lj} - q_{kj}q_{il}}{\sqrt{1-p}}. \tag{5}$$

Consequently,

$$\sum_{k,l} q_{kl}^2 (Q\Omega)_{kl}^2 = \frac{C_{ij} + p - 2p^2}{1-p}, \tag{6}$$

$$C_{ij} = (MM^\top M)_{ij} = \sum_{k,l} M_{il} M_{kl} M_{kj}.$$

It remains to compute the multiplier term in (3). This is the step where full stationarity is essential. Since $A = A^\top$,

$$Ax = A^\top Q^\top e_i = (Q^{\circ 3})^\top QQ^\top e_i = (Q^{\circ 3})^\top e_i = x^{\circ 3}.$$

It follows that

$$x^\top Ax = s_i, \qquad e_j^\top Ae_j = t_j, \qquad x^\top Ae_j = e_j^\top Ax = q^3.$$

Moreover, $\Omega^\top\Omega = xx^\top + vv^\top$. Therefore

$$\operatorname{tr}(A\Omega^\top\Omega) = \frac{s_i + t_j - 2p^2}{1-p}. \tag{7}$$

Combining (3), (6), and (7) proves the exact identity

$$\boxed{\mathcal{H}_Q(\Omega) = \frac{N_{ij}}{1 - M_{ij}}, \qquad N_{ij} = 3(MM^\top M)_{ij} + 3M_{ij} - 4M_{ij}^2 - s_i - t_j.} \tag{8}$$

6.2 Strict positivity at a largest squared entry

Assume first that every entry of M is strictly below one. Choose a largest entry

$$m = M_{ij} = \max_{k,l} M_{kl}.$$

Since M is nonempty and doubly stochastic, $0 < m < 1$. Write $S = s_i + t_j$. In

$$C_{ij} = \sum_{k,l} M_{il} M_{kl} M_{kj},$$

the slice $k = i$ contributes ms_i, the slice $l = j$ contributes mt_j, and their common term is m^3. Hence the following is an exact disjoint decomposition:

$$C_{ij} = mS - m^3 + R, \qquad R = \sum_{\substack{k \neq i \\ l \neq j}} M_{il} M_{kl} M_{kj} \geq 0. \tag{9}$$

The numerator in (8) becomes

$$N_{ij} = 3R + (3m - 1)S + 3m - 4m^2 - 3m^3. \tag{10}$$

There are two cases.

Case 1: $m \geq \frac{1}{3}$. The selected row and column each contain the entry m, so

$$S = 2m^2 + D_+, \qquad D_+ = \sum_{l \neq j} M_{il}^2 + \sum_{k \neq i} M_{kj}^2 \geq 0.$$

Substitution in (10) gives

$$N_{ij} = 3R + (3m - 1)D_+ + 3m(1 - m)^2 > 0. \tag{11}$$

Case 2: $m \leq \frac{1}{3}$. Maximality of m and double stochasticity give

$$s_i = \sum_l M_{il}^2 \leq m \sum_l M_{il} = m, \qquad t_j = \sum_k M_{kj}^2 \leq m \sum_k M_{kj} = m.$$

Thus

$$S = 2m - D_-, \qquad D_- = (m - s_i) + (m - t_j) \geq 0.$$

Now (10) gives

$$N_{ij} = 3R + (1 - 3m)D_- + m(1 - m)(1 + 3m) > 0. \tag{12}$$

At $m = 1/3$, the two formulas agree. Since $1 - m > 0$, (8)–(12) prove

$$\mathcal{H}_Q(\Omega) > 0.$$

More quantitatively, the displayed direction satisfies

$$\mathcal{H}_Q(\Omega) \geq \begin{cases} 3m(1-m), & m \geq 1/3, \\ m(1+3m), & m \leq 1/3. \end{cases} \tag{13}$$

6.3 Entries of magnitude one

The construction (4) is deliberately not used when $M_{ij} = 1$. In that case $q_{ij} = \pm 1$, and the unit norms of row i and column j force every other entry in that row and column to vanish. After row and column signed permutations,

$$Q = (1) \oplus Q', \qquad Q' \in \mathcal{O}(r-1).$$

The stationarity matrix is block diagonal, so stationarity of Q passes to Q'. Remove all such signed singleton blocks. If no block remains, the original Q was a signed permutation. Otherwise the remaining stationary orthogonal block has a largest squared entry $m \in (0, 1)$, so the construction above gives a strict ascent direction in that block. Extending its skew matrix by zero gives the same strict ascent direction for the original Q.

We have now proved that every stationary non-signed-permutation has an explicit Ω with $f_r''(0) = 4\mathcal{H}_Q(\Omega) > 0$. Since the first derivative vanishes at a stationary point, Taylor expansion gives

$$f_r(Qe^{t\Omega}) = f_r(Q) + 2\mathcal{H}_Q(\Omega)t^2 + o(t^2) > f_r(Q)$$

for all sufficiently small nonzero t. Such a point cannot be a local maximum.

6.4 Signed permutations are the strict global maxima

For every row of Q,

$$\sum_j q_{ij}^4 \leq \left(\sum_j q_{ij}^2 \right)^2 = 1.$$

Thus $f_r(Q) \leq r$. Equality holds precisely when every row has exactly one nonzero entry, of magnitude one; column orthogonality then makes Q a signed permutation.

At $Q = I$, one has $A = I$. Since every skew matrix has zero diagonal, the first term in (3) vanishes and

$$\mathcal{H}_I(\Omega) = -\operatorname{tr}(\Omega^\top \Omega) = -\|\Omega\|_F^2 < 0$$

for every nonzero skew Ω. Left and right multiplication by signed permutations preserves the objective and Hessian inertia, so every signed permutation is a strict local maximum for $r \geq 2$. For $r = 1$, the two points of $\mathcal{O}(1)$ are isolated and are both signed permutations. This completes the theorem.

7 Edge cases and degeneracies

The proof does not use a hidden genericity assumption.

(a) **$r = 1$.** $\mathcal{O}(1) = \{[-1], [1]\}$, and both points are signed permutations with value 1.

(b) **$r = 2$.** The general proof applies without modification. As a direct check, an orthogonal matrix is, up to signs and permutations, a planar rotation and has objective $2(\cos^4\theta + \sin^4\theta)$. Its local maxima occur at $\theta \in (\pi/2)\mathbb{Z}$, exactly the signed permutations. The 45° points have an ascent direction for maximization.

(c) **Zero entries.** The pivot construction never divides by q_{ij}. A largest entry of the doubly stochastic matrix M is positive, and the only denominator is $1 - m$, which is positive after signed singleton blocks are removed.

(d) **Repeated magnitudes.** The largest entry need not be unique. Any maximizing index pair satisfies the same decomposition and strict estimate.

(e) **Block diagonal and disconnected support.** An entry of magnitude one is handled by exact block peeling. More general orthogonal direct sums cause no problem: select a largest squared entry in a nontrivial residual block, build Ω there, and extend it by zero.

(f) **Hadamard-type stationary points.** If a normalized real Hadamard matrix of order $r > 1$ is used, then $M = J_r/r$, $s_i = t_j = 1/r$, and $C = M$. Formula (8) gives

$$\mathcal{H}_Q(\Omega) = \frac{4}{r} > 0.$$

Thus these highly repeated, full-support stationary points are explicitly excluded as local maxima.

(g) **Pairwise-stable but mixed-direction-unstable points.** The proof uses the full Hessian direction (4), not an assumption that elementary row or column rotations suffice. It therefore also handles structured points for which all coordinate Givens curvatures are nonpositive but a mixed direction has positive curvature.

(h) **Orthostochastic singularities.** Although $M = Q^{\circ 2}$ is doubly stochastic, the proof never replaces the orthostochastic image by the Birkhoff polytope. Double stochasticity is used only in the elementary inequalities for M; the ascent direction is lifted explicitly to the feasible curve $Qe^{t\Omega}$.

8 Proof scope and technical checks

For completeness, we record how the proof addresses the standard degeneracies that can arise in entrywise optimization on the orthogonal group.

1. **Square-only formulation:** every matrix is $r \times r$; no rectangular Stiefel problem is introduced.

2. **Full stationarity:** symmetry of $Q^\top Q^{\circ 3}$ is derived from all skew directions, rather than inferred from an incomplete test.

3. **Second-order sign and factor:** direct differentiation confirms $f''(0) = 4\mathcal{H}_Q(\Omega)$; positive curvature is ascent for maximization.
4. **Feasibility:** Ω in (4) is explicitly skew, and $Qe^{t\Omega} \in \mathcal{O}(r)$ for every real t.
5. **Stationarity usage:** the identity $Ax = x^{\circ 3}$ is invoked only after $A = A^\top$ has been established. It is not assumed for a generic orthogonal matrix.
6. **Index order:** the weighted term expands as $C_{ij} = (MM^\top M)_{ij} = \sum_{k,l} M_{il}M_{kl}M_{kj}$, ruling out a hidden transpose error.
7. **Zeros and repeated magnitudes:** no nonzero-entry or uniqueness assumption is made.
8. **Endpoint $m = 1$:** the undefined pivot is never used there; signed singleton blocks are removed exactly.
9. **Case split:** for $3m-1 \geq 0$ the lower bound $S \geq 2m^2$ is used, while for $3m-1 \leq 0$ the upper bound $S \leq 2m$ is used. The inequality direction is therefore correct in both branches.
10. **Strictness:** the residual terms $3m(1-m)^2$ and $m(1-m)(1+3m)$ are strictly positive for $0 < m < 1$; the boundary $m = 1/3$ has no gap.
11. **Blocks and degeneracies:** singleton blocks, arbitrary direct sums, normalized Hadamard points, and pairwise-insufficient examples are all covered by the same pivot mechanism.
12. **Local versus global:** the pivot proof rules out every non-permutation local maximum, while the separate rowwise inequality proves global optimality.
13. **Strict-saddle wording:** the precise conclusion is $\lambda_{\max}(\operatorname{Hess} f_r) > 0$ at each non-permutation stationary point; Hessian indefiniteness is not claimed.
14. **No circular classification or black box:** no classification of stationary points, external database, numerical search, or theorem of comparable difficulty is used in the proof.

References

[1] A. Edelman, T. A. Arias, and S. T. Smith, The geometry of algorithms with orthogonality constraints, *SIAM Journal on Matrix Analysis and Applications* **20** (1998), no. 2, 303–353. `https://doi.org/10.1137/S0895479895290954`

[2] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.

[3] T. Banica and I. Nechita, Almost Hadamard matrices: the case of arbitrary exponents, *Discrete Applied Mathematics* **161** (2013), no. 16–17, 2367–2379. `https://doi.org/10.1016/j.dam.2013.05.012`