The Jacobian Singularity in the Thomson Problem

Ramūnas Girdziušas* and Dario Ringach
Department of Neurobiology, UCLA.
(Dated: August 31, 2009)

The Thomson problem with the length constraint possesses a singular Jacobian matrix.

PACS numbers:

I. INTRODUCTION

Let $r_{i,j}$ denote the Euclidean distance between the points x_i and x_j in $\mathbb{R}^d$. Consider maximizing

$$E = \sum_{i<j} r_{i,j}^{-1}, \tag{1}$$

subject to $r_{i,i} = 1$ for $i = 1, \ldots, n$. This is the classical Thomson problem of locating n charges on a sphere.

An interesting modification occurs when the sum $\sum_{i<j}$ is replaced with $\sum_{i<j+1}$, and thus exluded distances are further constrained so that for $i = 1, \ldots, n$, with $n+1 \equiv 1$,

$$r_{i,i+1} = \frac{2\pi\alpha}{n},\ \alpha > 1. \tag{2}$$

The solutions then resemble the line-like arrangements.

Both problems are relevant to statistical modeling such as the covariance selection, coding, and the dimensionality reduction, see Ref. 1 for an overview.

It has been noticed by one of the authors of this note that the Thomson problem may serve as a generative model for the layer V1 neurons of the mammalian visual cortex. The point $\boldsymbol{x}_i$ would contain the normalized responses of the population of neurons which reacts to the stimulus bar oriented along the ith spatial direction.

The criterion E is nonpolynomial, and also nonlocal as all the index pairs (i, j) are connected in E. Little is known about the exact solution and how the length constraint may affect it, clf. Ref. 1 which provides numerical investigations and the Fourier analysis of the modified Thomson problem.

However, the criterion given by Eq. (1) could be replaced with the function which is polynomial w.r.t. to the variables r_{ij}. Indeed, such a criterion appears in the literature on the maximal entropy-based covariance selection, which is thoroughly discussed in Refs. 3–6, where Gaussian modeling reduces the entropy criterion to the maximization of the logarithm of the determinant of the (distance) matrix whose ijth element is the distance r_{ij}. The first-order optimality then demands setting to zero the elements in the inverse distance matrix whose index pair (i, j) corresponds to the unknown element r_{ij} in the original distance matrix. The earliest appearance of this idea is in Ref. 2!

The determinantal criterion is complicated, and no exact solutions are presently known. Moreover, the peculiarities of the distance-generative mechanism such as the demand for the Euclidean space with a fixed dimension, and the sphericity requirement, bring in another set of difficulties, e.g. the distance matrix may become singular.

Therefore, it seems reasonable to begin investigations of the Thomson problem by using the criteria which would include the polynomials whose degree is not higher than two. In principle, one may start applying the multivariate polynomial solvers to enumerate all possible solutions, but careful analysis is needed before the criteria could be formulated correctly.

In this note we shall discuss the difference between the optimal points of the original and modified Thomson problem in the case with the quadratic criteria. Our main result is encapsulated into the theorem given in Section II B, the precluding material introduces the criteria, and the remaining text states the proof.

II. MAIN RESULT

A. Formulation With Quadratic Criteria

Let us replace the criterion in Eq. (1) with

$$E = -\sum_{i<j} r_{i,j}^2, \tag{3}$$

and the n constraints in Eq. (2) with a single one:

$$\sum_{i=1}^{n} r_{i,i+1}^2 = \frac{(2\pi\alpha)^2}{n}. \tag{4}$$

Intuitively, the points should lie as far as possible from one another but the squared perimeter computed over their neigboring pairs, introduced by the ordering w.r.t the increasing index i, should remain fixed.

One can now gather all the unknown components of the points into the grand column-vector of length nd:

$$u = (x_1(1), \ldots, x_1(d), x_2(1), \ldots, x_n(d))'. \tag{5}$$

The quadratic Thomson problem (P1) now becomes:

$$u^* = \max_u u' Q_0 u, \tag{6}$$

$$\text{s.t. } u' Q_i u = 1,\ i = 1, \ldots, n. \tag{7}$$

*Electronic address: `ramunasg@ucla.edu`

The modified problem (P2) adds the constraint

$$u'Q_{n+1}u = l^2, \tag{8}$$

where l^2 is the constant given by the l.h.s. of Eq. (4). The symmetric matrices Q_i will be idetified in Section III A.

In order to discuss the optimality of the solutions of the problem P1 (or P2), one introduces n (or $n+1$) Lagrange multipliers $\lambda_i \in \mathbb{R}$, and studies the extrema of the Lagrangians L_1 and L_2:

$$L_k(u,\lambda) = \sum_{i=0}^{n+k-1} \lambda_i (u'Q_i u - p_i),\ k = 1,2. \tag{9}$$

Here $\lambda_0 = 1$, $p_0 = 0$, $p_i = 1$ for $i = 1,\ldots,n$, and $p_{n+1} = l^2$.

The local extrema $v^* = (u^*, \lambda^*)$ must then satify either $f_1(v^*) = 0$ (P1), or $f_2(v^*) = 0$ (P2), where

$$f_1(v) = \begin{pmatrix} A_1(\lambda)u \\ u'Q_1u - 1 \\ \vdots \\ u'Q_nu - 1 \end{pmatrix},\ f_2(v) = \begin{pmatrix} A_2(\lambda)u \\ u'Q_1u - 1 \\ \vdots \\ u'Q_nu - 1 \\ u'Q_{n+1}u - l^2 \end{pmatrix}, \tag{10}$$

and

$$A_k(\lambda) = 2\sum_{i=0}^{n+k-1} \lambda_i Q_i,\ k = 1,2. \tag{11}$$

Eqs. 10 present $nd+n$ (or $nd+n+1$) equations in the case P1 (or P2). In both cases the number of unknown quantities is equal to the number of equations.

B. The Jacobian Singularity

A further linearization of Eqs. (10) yields:

$$f_k(v) = v + J_k(v)v = 0,\ k = 1,2, \tag{12}$$

where the Jacobians J_k are:

$$J_1(u,\lambda) = \begin{pmatrix} A_1(\lambda) & B'(u) \\ B(u) & 0 \end{pmatrix}, \tag{13}$$

$$J_2(u,\lambda) = \begin{pmatrix} A_2(\lambda) & B'(u) & w \\ B(u) & 0 & 0 \\ w' & 0 & 0 \end{pmatrix}. \tag{14}$$

Here $w = Q_{n+1}u$ is an nd-dimensional column vector, and the matrix B is of the size $n \times nd$, whose ith row is given by

$$[B]_i = (0,\ldots,0,\underbrace{x_i(1),\ldots,x_i(d),}_{i\text{th block, } d \text{ elements}} 0,\ldots,0). \tag{15}$$

Despite the apparent similarity between the problems P1 and P2, the following result takes place.

Theorem. *Let u be a vector of the coordinates of the points as in Eq. (5), and assume that all the points are on a sphere in $\mathbb{R}^d$. In addition, let n real Lagrange multipliers satisfy*

$$\lambda_i \neq -\sum_{p=2}^{m} s_i(p), \tag{16}$$

where $i = 1,\ldots,n$, and the quantities $s_i(p)$ are given by Eq. (41).

The following is true:

1. *$J_1(u,\lambda)$ is nonsingular.*
2. *$J_2(u,\lambda)$ is singular for all values of the Lagrange multipliers.*

The proof will be split into separate sections which we begin by discussing the matrices Q_i.

III. PROOF

A. The Structure of Q_i

In order to identify the matrices in Eqs. (6)–(8), we shall rewrite the criteria in Eqs. (3) and (4) by introducing the permutation matrix

$$P_n = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1 \\ 1 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix}_{n\times n}. \tag{17}$$

In what follows a single subscript will ocassionally indicate the size of a square matrix, but when the matrix dimensions are clear, the subscript may also denote a different matrix within a certain family.

The reader may now verify that

$$Q_0 = \sum_{p=2}^{m} R_p' R_p,\quad Q_{n+1} = R_1' R_1, \tag{18}$$

where the upper limit for the sum arises because only a half of the cyclical shifts are needed to generate all the distinct pairs of the spatial points:

$$m = \begin{cases} n/2 & \text{if } n \text{ is even,} \\ (n-1)/2 & \text{if } n \text{ is odd.} \end{cases} \tag{19}$$

The matrices R_p turn out to be

$$R_p = (P_n^p - I_n) \otimes I_d. \tag{20}$$

The symbol $\otimes$ denotes the matrix Kronecker product. For $i = 1,\ldots,n$, the matrices Q_i are diagonal:

$$Q_i = \mathrm{diag}(0,\ldots,0,\underbrace{1,\ldots,1}_{i\text{th block}},0,\ldots,0), \tag{21}$$

where the ith block contains d unity elements.

This concludes the identification of all the matrices. The reader will notice that the matrix R_1 is the gradient operator acting along the discretized line induced by the increasing index i of the point x_i.

The matrix Q_{n+1} is the Laplacian on the line. The structure of the discrete Laplacian operator can be demonstrated by employing two well-known identities $(A \otimes B)' = A' \otimes B'$ and $A \otimes BC \otimes D = AC \otimes BD$, which result in:

$$Q_{n+1} = (P_n' P_n - P_n' - P_n + I_n) \otimes I_d. \tag{22}$$

A further identity $P_n' = P^{n-1}$ then brings in the doubled unity diagonal, yielding a familiar structure of the discrete Laplacian operator.

B. Rank-based Considerations

The following result will be utilized.

Lemma 1 (Ref. 7)**.** *Let A be nonsingular, and B have the size $m \times n$ and rank m. Then the matrix*

$$C = \begin{pmatrix} A & B' \\ B & 0 \end{pmatrix} \tag{23}$$

is nonsingular.

Proof. The matrix C is singular if $u'Cu = 0$ (or, which is the same, $Cu = 0$) for some $u \neq 0$. In the block-partitioned case given by Eq. (23), there should exist a vector $u = (w, v)$ such that

$$\begin{pmatrix} A & B' \\ B & 0 \end{pmatrix} \begin{pmatrix} w \\ v \end{pmatrix} = 0. \tag{24}$$

Eq. (24) demands that both, $Bw = 0$ and $Aw = 0$, which is possible only when $w = 0$ as A is nonsingular. The statement $w = 0$ with Eq. (24) demands that $A'v = 0$, which is possible only if $v = 0$ as A' has a full column rank. Therefore, $u'Cu = 0$ holds only when $u = 0$. □

Lemma 2. *Let the matrix B be given by Eq. (15), and assume that all the points x_i lie on a sphere in $\mathbb{R}^d$. Then rank$(B) = n$. Moreover, appending the row $w' = u'Q_{n+1}'$ to the matrix B does not alter its rank.*

Proof. The matrix B is already in a row-reduced echelon form whose ith pivot equals to $x_i(k)$ for $i = 1, \ldots, n$, where the index k indicates the smallest among integers $1, \ldots, d$ such that $x_i(k) \neq 0$. It is clear that there exist no point $x_i(k)$ whose all coodinates would be identical to zero. Hence, there are n nonzero pivots, and rank$(B) = n$.

In order to prove the second part, let us use the identity rank$(A') = $ rank(A), and consider the joint matrix of size $nd \times (n+1)$:

$$F = (B', w). \tag{25}$$

In order to save space, it is enough to describe the process for the first d rows:

$$d \text{ rows of } F = \begin{pmatrix} x_1(1) & 0 & \ldots & 0 & w_1 \\ x_1(2) & 0 & \ldots & 0 & w_2 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ x_1(d) & 0 & \ldots & 0 & w_d \end{pmatrix}, \tag{26}$$

as the rest proceeds independently in a similar manner.

For the sake of simplicity, let us assume that $x_1(1) \neq 0$, otherwise one should find $x_1(k) \neq 0$ and swap the rows 1 and k. After the elementary operations, the matrix in Eq. (26) becomes left-equivalent to

$$\text{reduced } d \text{ rows of } F = \begin{pmatrix} 1 & 0 & \ldots & 0 & z_1 \\ 0 & 0 & \ldots & 0 & z_2 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & z_d \end{pmatrix}, \tag{27}$$

where

$$z_1 = \frac{w_1}{x_1(1)}, \tag{28}$$

$$z_k = w_k - x_1(k)\frac{w_1}{x_1(1)}, \; k = 2, \ldots, d. \tag{29}$$

If the coordinates $x_1(k)$ and the elements w_k were free parameters, one would generally obtain $z_2 \neq 0$, and then would zero-out the remaining elements z_k for $k = 3, \ldots, d$. Thus, the element z_2 would become a nonzero pivot in a row-reduced echelon form, yielding rank$(F) = n + 1$.

However, due to the both, the sphericity of u, and the equation $w = Q_{n+1}u$, the reduced block of d rows has the property that $z_k = 0$ for $k = 2, \ldots, d$, or, to state it explicitly,

$$\begin{vmatrix} x_1(1) & x_1(k) \\ w_1 & w_k \end{vmatrix} = 0, \; k = 2, \ldots n. \tag{30}$$

Thus, the appended vector w does not contribute with any new pivots, hence rank$(F) = n$. □

Eqs. (30) may seem 'magic', but they show that the sphericity constraint includes more quadratic invariants than n constraints $\|x_i\|^2 = 1$.

Let us choose fifteen points on a sphere in a three dimensional Euclidean space. If we insert $w = Q_{n+1}u$, into Eq. (30), processing the first $d = 3$ rows of the matrix F generates two identities:

$$\begin{vmatrix} x_1(1) & x_1(2) \\ x_2(1) + x_5(1) & x_2(2) + x_5(2) \end{vmatrix} = 0, \tag{31}$$

and

$$\begin{vmatrix} x_1(1) & x_1(3) \\ x_2(1) + x_5(3) & x_2(3) + x_5(3) \end{vmatrix} = 0. \tag{32}$$

In general, we get $d - 1$ such relationships at each stage of processing d rows of the matrix F. Thus, the total number of the identities equals $n(d - 1)$.

C. The Spectrum of A_k

Recall that in addition to the rank of the matrix B, the other significant part of Lemma 1 is the nonsingularity of the matrix A in Eq. (23). Thus, what remains to be discussed is the singularity of the matrices A_1 and A_2 given by Eqs. (11).

Here the elementary transformations seem to be unwieldy, but one may employ the spectral method to state the conditions when the matrices are singular. It is known that the permutation P is unitarily-diagonalizable, clf. Ref. 8:

$$P = W^{*}\Omega W, \tag{33}$$

$$[W]_{ij} = \frac{1}{\sqrt{n}}\omega^{-(i-1)(j-1)},\ i,j = 1,\ldots,n. \tag{34}$$

$$\Omega = \mathrm{diag}(1,\omega,\ldots,\omega^{n-1}), \tag{35}$$

$$\omega = e^{\sqrt{-1}\frac{2\pi}{n}}. \tag{36}$$

Furthermore, by exploring the techniques presented in Ref. 8, one can show that the matrices R_p are unitarily-diagonalizable too:

$$UR_pU^{*} = (\Omega_n - I_n)\otimes I_d\,, \tag{37}$$

$$U^{*} = W_n^{*}\otimes W_d^{*}. \tag{38}$$

The identity $P_n^{pT} = P_n^{n-p}$ reveals that

$$UR_p'R_pU^{*} = S_p\otimes I_d, \tag{39}$$

$$S_p = (\Omega_n^{n-p} - I_n)(\Omega_n^{p} - I_n). \tag{40}$$

The entries of the diagonal matrices S_p are real. It is easy to see that $\omega^n = 1$, and $(\omega^k)^{(n-p)} = \omega^{kn}\omega^{-kp} = \omega^{-kp}$, which leads to $\Omega^n = 1$ and $\Omega^{n-p} = \Omega^{-p}$, respectively. Therefore, the ith diagonal entry of the matrix S_p is:

$$s_i(p) = 2\left(1-\cos\frac{2\pi p(i-1)}{n}\right),\ i = 1,\ldots,n. \tag{41}$$

Notably, $s_1(p) = 0$, hence all the matrices S_p are singular. In turn, all the matrices $R_p'R_p$, and hence the matrices Q_0 and Q_{n+1} are singular too. However, various weighted sums of such matrices will remain singular only in exceptional cases.

In order to reveal the exceptional values of the Lagrange multipliers that may produce singular matrices A_1 and A_2, one may notice that the sums in Eqs. (11), performed over the index values $i = 1,\ldots,n+1$, can also be rewritten according to

$$\sum_{i=1}^{n}\lambda_i Q_i = \Lambda_n\otimes I_d, \tag{42}$$

$$\Lambda_n = \mathrm{diag}(\lambda_1,\ldots,\lambda_n). \tag{43}$$

The spectrum of A_2 becomes clear:

$$UA_2U^{*} = 2\left(\sum_{p=2}^{m}S_p + \Lambda_n + \lambda_{n+1}S_1\right)\otimes I_d, \tag{44}$$

$$= 2\,\mathrm{diag}(\underbrace{g_1,\ldots,g_1}_{\text{d times}},\ldots,\underbrace{g_n,\ldots,g_n}_{\text{d times}}), \tag{45}$$

$$g_i = \sum_{p=2}^{m}s_i(p) + \lambda_i + \lambda_{n+1}s_i(1). \tag{46}$$

The spectrum of the matrix A_1 is obtained by setting $\lambda_{n+1} = 0$.

D. Summary

The results contained in Lemmas 1 and 2, as well as Eqs. (44) and (46), lead to the theorem stated in Sect. II B.

IV. CONCLUSION

The Thomson problem with the quadratic criteria has an interesting phenomenon. Its Jacobian matrix, a central quantity that characterizes the first-order optimality of the solutions, becomes singular if one adds the constraint on the sum of the squared distances between the neighboring points on a sphere. Remarkably, this occurs not because of the singularity of the matrices in the quadratic forms which specify the maximal-distance criterion and the length constraint. The Jacobian singularity occurs because the length constraint depends linearly on the sphericity requirement after the linearization. The precise conditions for the phenomenon to take place have been derived.

[1] S. Alben, Phys. Rev. E. **78**, 066603 (2008).

[2] M. Fiedler, Numerische Math. **9**, 109 (1966).

[3] R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz, Lin. Alg. Appl. **58**, 109 (1984).

[4] H. Lev-Ari, S. R. Parker, and T. Kailath, IEEE Tr. Inf. Th. **35**, 497 (1989).

[5] H. Dym and I. Gohberg, Lin. Alg. Appl. **36**, 1 (1981).

[6] A. P. Dempster, Biometrics **28**, 157 (1972).

[7] J. Nocedal and S. Wright, *Numerical Optimization* (Springer, 1999).

[8] P. J. Davis, *Circulant Matrices* (Wiley-Interscience, 1979).