Memorandum 2034 (February 2014). ISSN 1874-4850. Available from: http://www.math.utwente.nl/publications
Department of Applied Mathematics, University of Twente, Enschede, The Netherlands

Quadratic maximization on the unit simplex: structure, stability, genericity and application in biology

Georg Still *   Faizan Ahmed†

December 25, 2013

Abstract

The paper deals with the simple but important problem of maximizing a (nonconvex) quadratic function on the unit simplex. This program is directly related to the concept of evolutionarily stable strategies (ESS) in biology. We discuss this relation and study optimality conditions, stability and generic properties of the problem. We also consider a vector iteration algorithm to compute (local) maximizers. We compare the maximization on the unit simplex with the easier problem of the maximization of a quadratic function on the unit ball.

1 Introduction

In the present paper we study the maximization of a (in general nonconvex) quadratic function on the unit ball and the unit simplex:

$$P_B: \qquad \max \frac{1}{2}x^T Ax \quad \text{st.} \quad x \in B_m := \{x \in \mathbb{R}^m \mid x^T x = 1\}$$

$$P_S: \qquad \max \frac{1}{2}x^T Ax \quad \text{st.} \quad x \in \Delta_m := \{x \in \mathbb{R}^m \mid e^T x = 1, x \geq 0\}$$

where $A = (a_{ij})$ is a symmetric $m \times m$-matrix, $e \in \mathbb{R}^m$ denotes the vector with all one's. Since we do not assume A to be positive semidefinite, both programs are nonconvex problems. However the global maximizer for P_B is polynomially (approximately) computable whereas the (global) maximization of P_S is NP-hard. In the paper we will shortly compare both programs P_B, P_S, and consider two similar vector iteration methods for computing a global solution of P_B and a (local) solution of P_S. Then, we study P_S in more detail. We present an application in evolutionary biology and analyze the structure, stability and generic properties of P_S.

The paper is organized as follows. In Section 2 we present two well-known vector iterations for solving the programs, and discuss convergence and monotonicity properties. Section 3 shortly introduces the concept of evolutionarily stable strategies (ESS) and studies the direct relation with P_S. We also present an example showing that the number of ESS's (strict local maximizers) of P_S may grow exponentially with the dimension m of the problem. In Section 4 we recall the optimality conditions for P_S also in terms of the ESS model. In Section 5 we apply results from parametric optimization to analyze the stability of the program P_S wrt. small perturbations of the matrix A.

*Corresponding author: University of Twente, Department of Mathematics, Email: g.j.still@utwente.nl
†University of Twente, Department of Mathematics, Email: f.ahmed@utwente.nl

Section 6 deals with genericity results concerning P_S and P_B.

We tried to present the topic in such a form that it might be interesting for both, scientists in biology and in optimization.

Throughout the paper, for $x \in \mathbb{R}^m$, $\|x\|$ denotes the Euclidean norm and $\mathcal{N}_\varepsilon(\overline{x}) = \{x \in \mathbb{R}^m \mid \|x - \overline{x}\| \leq \varepsilon\}$ is the ε-neighborhood of $\overline{x} \in \mathbb{R}^m$. Furthermore, $\mathcal{S}_m$ denotes the set of symmetric $(m \times m)$-matrices and for $A \in \mathcal{S}_m$, by $\|A\|$ we mean the Frobenius norm, $\|A\| = (\sum_{ij} a_{ij})^{1/2}$.

2 Vector iteration for solving the problems

It is well-known that the global maximizers of P_B are precisely the (normalized) eigenvectors corresponding to the largest eigenvalue λ_1 of A. So, by replacing A by $A + \alpha I$, with α large enough, we can assume wlog. that A is positive definite (in fact we can chose $\alpha > -\lambda_m$ where λ_m is the smallest eigenvalue of A). Similarly by defining the matrix $E := [e, \ldots, e] \in \mathcal{S}_m$ (all one's) the local maximizers of P_S wrt. A and wrt. $A + \alpha E$, $(\alpha \in \mathbb{R})$ coincide. Indeed, by noticing that $x \in \Delta_m$ satisfies $x^T E x = e^T x = 1$, we obtain for $x, y \in \Delta_m$:

$$x^T(A + \alpha E)x \geq y^T(A + \alpha E)y \quad \Leftrightarrow \quad x^T A x + \alpha \geq y^T A y + \alpha \quad \Leftrightarrow \quad x^T A x \geq y^T A y$$

So, in P_S wlog. we can assume $A > 0$, *i.e.*, $a_{ij} > 0, \forall i, j$. Let us now consider the following vector iterations:

For P_B : Starting with $x_0 \in B_m$ iterate

$$x_{k+1} = \frac{Ax_k}{\|Ax_k\|}, \quad k = 0, 1, \ldots \tag{Iter$_B$}$$

For P_S: Start with $x_0 \in \Delta_m$ and iterate for $k = 0, 1, \ldots$,

$$x_{k+1} \in \Delta_m \text{ is defined by:} \quad [x_{k+1}]_i = \frac{[x_k]_i \cdot [Ax_k]_i}{x_k^T A x_k}, \quad i = 1, \ldots, m\ . \tag{Iter$_S$}$$

Here, $[x_k]_i$ denotes the ith component of $x_k \in \mathbb{R}^m$. The next theorem describes the convergence and monotonicity properties of these iterations.

Theorem 1. *[*convergence and monotonicity results*]*

(1) Let $A \in \mathcal{S}_m$ be a positive definite matrix with eigenvalues $\lambda_1 > \lambda_2 > \ldots > 0$ and eigenspace S_1 corresponding to the largest eigenvalue λ_1. We assume that the starting vector x_0 is not orthogonal to S_1. Then, for Iter$_B$ the following holds.

(a) For the distance dist $(x_k, S_1) := \min\{\|y - x_k\| \mid y \in S_1 \cap B_m\}$ *between x_k and S_1,*

$$\text{dist}\,(x_k, S_1) = \mathcal{O}\left(\tfrac{\lambda_2}{\lambda_1}\right)^k .$$

(b) The Rayleigh quotients $x_k^T A x_k$ satisfy the monotonicity property,

$$x_k^T A x_k \leq x_{k+1}^T A x_{k+1}\ .$$

(2) Let be given a matrix $A \in \mathcal{S}_m$, $A > 0$. Then, also for Iter$_S$ the monotonicity holds: $x_k^T A x_k \leq x_{k+1}^T A x_{k+1}$.

Proof: For the convergence rate in (1) we refer to [8, Sect. 7.3]. The monotonicity property (1)(b) is proven in an unpublished note [9].
We add the proof: For $x_0 \neq 0$ we define $a_\nu := x_0^T A^\nu x_0$ for $\nu = 1, \dots, 3$ and note that these numbers are positive (as inner product $y^T y$, or quadratic form with positive definite A). We now show

$$\rho_0 \le \rho_1 \quad \text{for} \quad \rho_0 := \frac{x_0^T A x_0}{x_0^T x_0} = \frac{a_1}{a_0} \text{ and } \rho_1 := \frac{x_0^T A^3 x_0}{x_0^T A^2 x_0} = \frac{a_3}{a_2} \, .$$

For the number $Q_1 := (a_3 x_0 - a_2 A x_0)^T A (a_3 x_0 - a_2 A x_0) \ge 0$ (A is positive definite) we obtain

$$\begin{aligned} 0 &\le Q_1 = (a_3 A x_0 - a_2 A^2 x_0)^T (a_3 x_0 - a_2 A x_0) \\ &= a_3^2 a_1 - 2 a_3 a_2 a_2 + a_2^2 a_3 = a_3^2 a_1 - a_3 a_2^2 \end{aligned}$$

and after division by $a_1 a_2 a_3$ we find

$$\frac{a_3}{a_2} - \frac{a_2}{a_1} \ge 0 \quad \text{or} \quad \frac{a_3}{a_2} \ge \frac{a_2}{a_1} \, . \tag{1}$$

Similarly,

$$0 \le (a_2 x_0 - a_1 A x_0)^T (a_2 x_0 - a_1 A x_0) = a_2^2 a_0 - 2 a_2 a_1^2 + a_1^2 a_2 = a_2^2 a_0 - a_2 a_1^2$$

and after division by $a_0 a_1 a_2$ we find $\frac{a_2}{a_1} - \frac{a_1}{a_0} \ge 0$ or $\frac{a_2}{a_1} \ge \frac{a_1}{a_0}$. Together with (1) this yields $\frac{a_3}{a_2} \ge \frac{a_1}{a_0}$ and so, $\rho_0 \le \rho_1$. The monotonicity in (2) has been shown in [12].

□

According to the preceding theorem (under mild assumptions), the iterate x_k in Iter_B converges linearly to the eigenspace S_1, *i.e.*, to the set of global maximizers of P_B. For Iter_S it can only be expected that x_k converges to a local maximizer (or a fixed point of Iter_S). The global convergence behavior is more complicated (see *e.g.*, [3] for details).

3 Evolutionarily stable strategies in biology

In this section we discuss a model in evolutionary biology. We introduce the concept of an evolutionarily stable strategy and deal with its direct relation with the program P_S. We emphasize that in our paper we restrict the discussion to symmetric matrices.

According to Maynard Smith [14] we consider a population of individuals which differ in m distinct features (also called strategies or genes) as follows:

- For $x = (x_1, \dots, x_m) \in \Delta_m$, the component x_i gives the percentage of the population with feature i. So, x gives the strategy (state) of the whole population.
- We have given a symmetric fitness matrix $A = (a_{ij}) > 0$. The elements $a_{ij} > 0$ can be seen as the fitness factor for feature i combined with feature j. A large value a_{ij} means that a combination of features j and i in the population contributes largely (with factor $a_{ij} x_i x_j$) to the fitness of the population.
- The value $x^T A x$ then gives the (mean) fitness of a population with strategy x.

In the model it is assumed that the fitness increases leading to

Definition A [ESS] *Given a fitness matrix $A \in \mathcal{S}_m$, the vector $x \in \Delta_m$ is called evolutionarily stable strategy (ESS) for A if there is some $\alpha > 0$ such that*

$$\big(x + \rho(y-x)\big)^T A\big(x + \rho(y-x)\big) \;<\; x^T Ax \quad \forall x \neq y \in \Delta_m,\; 0 < \rho \leq \alpha. \tag{2}$$

In words: any perturbation $x + \rho(y - x)$ of the population with strategy x by a small group of individuals with strategy y is not profitable.

By noticing that a neighborhood of $x \in \Delta_m$ given by $N^1_\alpha = \{x + \rho(y - x) \mid y \in \Delta_m, 0 < \rho \leq \alpha\}, \alpha > 0$ contains a (common) neighborhood

$$N_\varepsilon(x) = \{y \in \Delta_m \mid \|y - x\| \leq \varepsilon\}, \varepsilon > 0$$

and vice versa, with the standard definition for a (strict) local maximizer we directly conclude

Lemma 1. *Let be given $A \in \mathcal{S}_m$ and $x \in \Delta_m$. Then, x is an ESS for A if and only if x is a strict local maximizer of P_S wrt. A.*

In evolutionary biology commonly another (equivalent) definition for ESS is used. To obtain this, we write (2) equivalently (after dividing by $\rho > 0$) as:

$$\rho(y-x)^T A(y-x) + 2(y-x)^T Ax \;<\; 0 \quad \forall x \neq y \in \Delta_m,\; 0 < \rho \leq \alpha\;.$$

This condition is obviously equivalent with

$$(y-x)^T Ax \leq 0, \quad \text{and in case of equality we have } (y-x)^T A(y-x) < 0\;,$$

which can be re-written as

Definition B [definition of ESS in biology] *A point $x \in \Delta_m$ is called an ESS for A if we have:*

(1) $y^T Ax \leq x^T Ax \;\; \forall y \in \Delta_m$ and

(2) if $y^T Ax = x^T Ax$ holds for $x \neq y \in \Delta_m$ then $\;y^T Ay < y^T Ax$.

We shortly discuss the interesting question of how much ESS (*i.e., strict local maximizers*) a matrix $A \in \mathcal{S}_m$ may possess. It has been shown in [5] that the number of ESS of $A \in \mathcal{S}_m$ can grow exponentially with m. As a concrete example (obtained by the construction in [5]) consider for $m = 3 \cdot k$, $k \in \mathbb{N}$ the matrix

$$A = \begin{pmatrix} I & C & \dots & C \\ C & I & \dots & C \\ \vdots & \vdots & \ddots & \vdots \\ C & \dots & C & I \end{pmatrix} \in \mathcal{S}_m \quad \text{with} \quad C := \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$$

and I the (3×3)-unit matrix. It is not difficult to see that this matrix has $3^k = (3^{1/3})^m$ different ESS (isolated, global maximizers). More precisely, for any choice of an index set $J = \{i_1, \dots, i_k\}$ with $i_j \in \{1, 2, 3\}$ (3^k possibilities), we define the coefficients of an vector $x = x(J) \in \Delta_m$ as follows:

$$x_i = \frac{1}{k} \quad \text{if} \quad i = 3(j-1) + i_j,\; j = 1, \dots, k\;, \quad \text{and} \quad x_i = 0,\; i \text{ otherwise.}$$

Then each such $x = x(J)$ yields an ESS with the same maximum value $x^T Ax = 2 - \frac{1}{k}$. The fact that the number of strict local maximizer of P_S can grow exponentially with m "indicates" that the problem is NP-hard. For a formal NP-hardness proof we refer to [13].

4 Optimality conditions

In this section we present optimality conditions for P_S in the context of optimization and evolutionary biology. Some of these results will be used in the stability analysis of Section 5.

In optimization, optimality conditions are usually given in terms of the Karush-Kuhn-Tucker condition (KKT). To do so, we introduce the index set $M := \{i = 1, \dots, m\}$ and recall the program P_S with $A \in \mathcal{S}_m$:

$$P_S : \qquad \max \frac{1}{2} x^T A x \quad \text{st.} \quad x \in \Delta_m := \{x \in \mathbb{R}^m \mid e^T x = 1, \;\; x_i \geq 0, \forall i \in M\}$$

As usual, we define the active index set with respect to the constraints $x_i \geq 0$, $I(x) := \{i \in M \mid x_i = 0\}$. For a point $x \in \Delta_m$ the KKT condition is said to hold if there exist Lagrange-multipliers $\lambda \in \mathbb{R}$ and $\mu_i \geq 0, i \in M$, corresponding to the constraints $e^T x = 1$ and $x_i \geq 0$, such that

$$Ax - \lambda e + \sum_{i \in M} \mu_i e_i = 0, \;\; \text{and} \quad \mu_i x_i = 0, \; \forall i \in M \; . \tag{3}$$

Here, $e_i, \; i \in M$ denote the standart basis vectors in $\mathbb{R}^m$. Since for $x \in \Delta_m$ in particular $x \neq 0$ holds, not all constraints $e^T x = 1, x_i \geq 0, i \in M$, can be active simultaneously. Thus, the active gradients $e, e_i, i \in I(x)$ are always linearly independent. So, the linear independency constraint qualification (LICQ) is automatically fulfilled at any feasible point $x \in \Delta_m$.

Hence, according to standard results in optimization, for any local maximizer of P_S the KKT condition must hold with unique multipliers λ, μ (unique by LICQ) (see *e.g.*, [7, Th. 21.7]).

Strict complementarity is said to hold at a solution (x, λ, μ) of (3) if we have:

$$\mu_i > 0 \quad \text{for all} \;\; i \in I(x) \; . \tag{SC}$$

In the context of evolutionary biology, necessary optimality conditions are usually formulated in terms of the following index sets. For a point $x \in \Delta_m$ we define

$$R(x) := \{i \in M \mid x_i > 0\} \quad \text{and} \quad S(x) := \{i \in M \mid [Ax]_i = \max_j [Ax]_j\} \tag{4}$$

If we write the KKT conditions componentwise,

$$[Ax]_i = \lambda - \mu_i, \quad \mu_i = 0, i \in R(x), \;\; \mu_i \geq 0, i \in M \setminus R(x), \tag{5}$$

we see that $\lambda = [Ax]_i = \max_j [Ax]_j = x^T A x$, $i \in R(x)$, holds. So the KKT condition implies $R(x) \subset S(x)$ and from (5) we conclude the converse. Moreover, obviously, the condition SC is equivalent to $R(x) = S(x)$. Note also, that for $x \in \Delta_m$ the relation $R(x) \subset S(x)$ implies with $\lambda := \max_j [Ax]_j$ (see (5) and Definition B(1)),

$$x^T A x = \sum_i x_i [Ax]_i = \sum_i x_i \lambda = \lambda = \sum_i y_i \lambda \geq y^T A x \quad \text{ for any } y \in \Delta_m \; .$$

Also the converse is true. Summarizing we obtain.

Lemma 2. *Given $A \in \mathcal{S}_m$, the following are equivalent necessary conditions for $\overline{x} \in \Delta_m$ to be a local maximizer of P_S:*

$$\text{the KKT condition (3) holds} \;\; \Leftrightarrow \;\; R(\overline{x}) \subset S(\overline{x}) \;\; \Leftrightarrow \;\; \overline{x}^T A \overline{x} \geq y^T A \overline{x}, \; \forall y \in \Delta_m \;\; \text{holds} \; .$$

Moreover, $\overline{x}$ satisfies the KKT condition with SC iff $R(\overline{x}) = S(\overline{x})$.

Since P_S is not convex (in general, A may be indefinite) the KKT condition (cf. Lemma 2) need not be sufficient for optimality and second order conditions are needed. To do so, as usual for a KKT point $\overline{x}$ we have to consider the cone of "critical directions",

$$C_{\overline{x}} = \{d \in \mathbb{R}^m \mid d^T A\overline{x} \geq 0, e^T d = 0, e_i^T d \geq 0,\ i \in I(\overline{x})\}\ .$$

By using the KKT condition, this cone simplifies to

$$C_{\overline{x}} = \{d \in \mathbb{R}^m \mid e^T d = 0;\ \ e_i^T d = 0,\ \text{if } \mu_i > 0;\ \ e_i^T d \geq 0,\ \ \text{if } \mu_i = 0, i \in I(\overline{x})\}\ .$$

Note that the program P_S has only linear constraints and a quadratic objective. Therefore, no higher order effects can occur so that in the second order conditions there is no gap between the necessary and sufficient part.

Lemma 3. *Let $A \in \mathcal{S}_m$. Then a point $\overline{x} \in \Delta_m$ is a strict local maximizer of P_S if and only if the KKT condition holds with second order condition:*

$$d^T Ad < 0 \quad \forall 0 \neq d \in C_{\overline{x}}\ . \tag{SOC}$$

The KKT point $\overline{x}$ is a local maximizer iff (the weak inequality) $d^T Ad \leq 0\ \ \forall 0 \neq d \in C_{\overline{x}}$ holds. Moreover, for a strict local maximizer the following growth condition (maximizer of order 2) is valid with some constants $\varepsilon, c > 0$,

$$\overline{x}^T A\overline{x} \geq x^T Ax + c\|x - \overline{x}\|^2 \qquad \forall x \in \Delta_m,\ \ \|x - \overline{x}\| \leq \varepsilon. \tag{6}$$

Proof: For the direction "⇐" of the optimality conditions see, *.e.g*, [7, Theorem 12.6]. An easy modification of the proof yields (6), *i.e.*, $\overline{x}$ is a so-called local maximizer of order two.
"⇒": We only show the strict maximizer case. Suppose to the contrary there is some $0 \neq d \in C_{\overline{x}}$ such that $d^T Ad \geq 0$. Then for small $\lambda > 0$ the vectors $\overline{x} + \lambda d$ are in Δ_m and using the KKT condition we find $d^T A\overline{x} = \lambda d^T e + \sum_{i \in I(\overline{x})} \mu_i d_i = 0$ and then

$$(\overline{x} + \lambda d)^T A(\overline{x} + \lambda d) = \overline{x}^T A\overline{x} + 2\lambda d^T A\overline{x} + \lambda^2 d^T Ad \geq \overline{x}^T A\overline{x}$$

contradicting the assumption that $\overline{x}$ is strict local maximizer. The weak case is similar.

□

In the stability analysis of the next section the following (extended) tangent space for a KKT point $\overline{x}$ will play an important role:

$$\begin{aligned} T^+_{\overline{x}} &= \{d \in \mathbb{R}^m \mid e^T d = 0,\ \ e_i^T d = 0\ \text{ if } \mu_i > 0,\ i \in I(\overline{x})\} \\ &= \{d \in \mathbb{R}^m \mid e^T d = 0,\ \ d_i = 0,\ i \in M \setminus S(\overline{x})\} \end{aligned} \tag{7}$$

We directly see that for an KKT point $\overline{x}$ we have:

$$C_{\overline{x}} \subset T^+_{\overline{x}} \quad \text{and} \quad C_{\overline{x}} = T^+_{\overline{x}} \ \text{ holds iff } \ R(\overline{x}) = S(\overline{x})\ . \tag{8}$$

For later purposes we add a lemma.

Lemma 4. *Let $\overline{x} \in \Delta_m$ be a local maximizer of P_S wrt. A with $R(\overline{x}) = S(\overline{x})$. If $\overline{x}$ is not a strict local maximizer we have $\det(A_{R(\overline{x})}) = 0$, where $A_{R(\overline{x})}$ denotes the principal submatrix of A corresponding to the index set $R(\overline{x})$.*

Proof: Recall that $R(\overline{x}) = S(\overline{x})$ implies $C_{\overline{x}} = T^{+}_{\overline{x}} = \{d \in \mathbb{R}^m \mid e^T d = 0, d_i = 0,\ i \in M \setminus R(\overline{x})\}$. By Lemma 3 the KKT condition must hold with SOC, implying $d^T A\overline{x} = 0$ and $d^T Ad \le 0 \ \ \forall d \in T^{+}_{\overline{x}}$. Since $\overline{x}$ is a nonstrict maximizer there must exist $0 \ne z \in T^{+}_{\overline{x}}$ such that $z^T Az = 0$. By defining $R := R(\overline{x})$ and $d_R := (d_i, i \in R)$ we thus have a vector $0 \ne z_R \in \mathbb{R}^{|R|},\ e_R^T z_R$ such that

$$\overline{x}_R A_R z_R = 0, \quad z_R^T A_R z_R = 0, \quad \text{and} \quad d_R^T A_R d_R \le 0 \quad \forall d_R \in \mathbb{R}^{|R|} \text{ with } e_R^T d_R = 0 \ . \qquad (9)$$

So, for any $\delta > 0$ in view of $e_R^T(z_R \pm \delta d_R) = 0$, for all d_R with $e_R^T d_R = 0$, we find

$$(z_R \pm \delta d_R)^T A_R (z_R \pm \delta d_R) = \delta^2 d_R^T A_R d_R \pm 2\delta d_R^T A_R z_R \le 0 \ .$$

By division by $\delta > 0$ and letting $\delta \downarrow 0$ it follows $d_R^T A_R z_R = 0$ for all $e_R^T d_R = 0$. Consequently, together with $\overline{x}_R A_R z_R = 0$ and $e_R^T \overline{x}_R = 1$ the vector $A_R z_R$ is perpendicular to a basis of $\mathbb{R}^{|R|}$ and thus $A_R z_R = 0$ must hold implying $\det(A_R) = 0$. XF □

5 Stability of an ESS

In this section we study the problem $P_S(A)$ in dependence of the matrix $A \in \mathcal{S}_m$ as a parameter:

$$P_S(A): \quad \max \ \frac{1}{2} x^T Ax \quad \text{st.} \quad x \in \Delta_m := \{x \in \mathbb{R}^m \mid e^T x = 1, x \ge 0\}$$

Let be given a matrix $\overline{A} \in \mathcal{S}_m$ and a strict local maximizer $\overline{x}$ of $P_S(\overline{A})$, *i.e.*, an ESS wrt. $\overline{A}$. We wish to know what may happen with the ESS $\overline{x}$ if the matrix $\overline{A}$ is slightly perturbed. How changes the ESS and may he possibly get lost? Such questions are studied in the field of Parametric Optimization (see *e.g.*, [6, 4]).

Our program $P_S(A)$ is especially easy, since the feasible set does not change. Only by using simple continuity arguments it can easily be seen that for $A \approx \overline{A}$ there must remain a local maximizer $x(A) \approx \overline{x}$ (at least one). However, the strict local maximizer can change into a nonstrict (nonunique) local maximizer, *i.e.*, the ESS $\overline{x}$ is lost. Such stability results have been proven in [2, Theorem 16] under the assumption $R(\overline{x}) = S(\overline{x})$.

By applying results from parametric optimization we can however give much preciser stability results. We start with a general Lipschitz stability statement (see also [4, Prop. 4.36] for a more general result).

Lemma 5. *Let $\overline{x}$ be a strict local maximizer of $P_S(\overline{A})$. Then there exist numbers $\varepsilon, \delta, L > 0$ such that for any $A \in \mathcal{N}\varepsilon(\overline{A})$ there exists a local maximizer $x(A) \in \mathcal{N}_\delta(\overline{x})$ (at least one) and for each such local maximizer $x(A)$ we have*

$$\|x(A) - \overline{x}\| \le L\|A - \overline{A}\| \ .$$

Proof: We firstly show the existence of (at least) one local maximizer $x(A)$ of A near $\overline{x}$. By putting $q(A, x) := x^T Ax/2$ and recalling that $\overline{x}$ is a strict local maximizer with max-value $m := q(\overline{A}, \overline{x})$, by continuity, there exist numbers $\varepsilon, \alpha, \delta > 0$ such that: (1) $q(A, \overline{x}) \ge m - \frac{\alpha}{2} \quad \forall A \in \mathcal{N}_\varepsilon(\overline{A})$, (2) $q(\overline{A}, x) \le m - 2\alpha \quad \forall x \in \Delta_m, \|x - \overline{x}\| = \delta$ and (3) $q(A, x) \le m - \alpha \quad \forall x \in \Delta_m, \|x - \overline{x}\| = \delta$ and all $A \in \mathcal{N}_\varepsilon(\overline{A})$. The existence of a local maximizer $x(A)$ with $\|x(A) - \overline{x}\| < \delta$ for $A \in \mathcal{N}_\varepsilon(\overline{A})$

follows from (1) and (3). Since $\overline{x}$ is a strict local maximizer of order 2 (see (6)) with some $c > 0, \delta > 0$ it holds:

$$q(\overline{A}, \overline{x}) - q(\overline{A}, x) \geq c\|x - \overline{x}\|^2 \quad \forall x \in \mathcal{N}_\delta(\overline{x}) \cap \Delta_m \; . \tag{10}$$

For a local maximizer $x := x(A) \in \mathcal{N}_\delta(\overline{x})$ we find $q(A, \overline{x}) - q(A, x) \leq 0$ and then

$$\begin{aligned} q(\overline{A}, \overline{x}) - q(\overline{A}, x) &= [q(A, x) - q(\overline{A}, x)] - [q(A, \overline{x}) - q(\overline{A}, \overline{x})] + [q(A, \overline{x}) - q(A, x)] \\ &\leq [q(A, x) - q(\overline{A}, x)] - [q(A, \overline{x}) - q(\overline{A}, \overline{x})] \\ &= \nabla_x \big[q(A, \overline{x} + \tau(x - \overline{x})) - q(\overline{A}, \overline{x} + \tau(x - \overline{x}))\big]^T (x - \overline{x}) \end{aligned}$$

with some $0 < \tau < 1$. In the last inequality we have applied the mean value theorem wrt. x for the function $q(A, x) - q(\overline{A}, x) = \frac{1}{2} x^T (A - \overline{A}) x$. By using $\nabla_x [q(A, x) - q(\overline{A}, x)] = (A - \overline{A}) x$ we find

$$q(\overline{A}, \overline{x}) - q(\overline{A}, x) \quad \leq \quad \max_{z \in \mathcal{N}_\delta(\overline{x})} \|A - \overline{A}\| \|z\| \|x - \overline{x}\|$$

Letting $\gamma := \max_{z \in \mathcal{N}_\delta(\overline{x})} \|z\|$, with (10) we obtain $c\|x - \overline{x}\|^2 \leq \gamma \|A - \overline{A}\| \cdot \|x - \overline{x}\|$ and the Lipschitz continuity result is valid with $L := \gamma / c$. □

We give an example where the ESS gets lost.

Example 1. The matrix $\overline{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ has the strict local maximizer $\overline{x} = (1, 0, 0)$. It is not difficult to see that for small $\alpha > 0$ the perturbed matrix

$$A_\alpha = \begin{pmatrix} 1 - 2\alpha & 1 - \alpha & 1 - \alpha \\ 1 - \alpha & 0 & -\alpha \\ 1 - \alpha & -\alpha & 0 \end{pmatrix}$$

has the nonstrict local maximizers $x_\rho = (1 - \rho)(1 - \alpha, \alpha, 0) + \rho(1 - \alpha, 0, \alpha), \; \rho \in [0, 1]$.

So, locally the ESS $\overline{x}$ is lost. Note that in this example we have $R(\overline{x}) = \{1\}, S(\overline{x}) = \{1, 2, 3\}$ and consequently, SC, *i.e.*, $R(\overline{x}) = S(\overline{x})$, is not fulfilled.

Recall, that in the preceding (bad) example the condition SC is not fulfilled. The next theorem shows that under SC strong stability holds. This result is a special case of a more general result (stability of so-called nondegenerate local maximizers in nonlinear optimization). The result goes back to Fiacco [6]. For completeness, we give a proof for our special program.

Theorem 2. *Let $\overline{x}$ be an ESS (strict local maximizer) of $\overline{A} \in \mathcal{S}_m$ with $R(\overline{x}) = S(\overline{x})$, i.e., the KKT condition holds with SC and the (strong) second order condition SOC.*

Then, there exist $\varepsilon, \delta > 0$ and a C^∞ (rational) function $x : \mathcal{N}_\varepsilon(\overline{A}) \to \mathcal{N}_\delta(\overline{x})$, $A \to x(A)$ with $x(\overline{A}) = \overline{x}$ and for any $A \in \mathcal{N}_\varepsilon(\overline{A})$ the vector $x(A)$ is an ESS of A and it is the unique local maximizer of A in $\mathcal{N}_\delta(\overline{x})$.

Proof: Let us define $\overline{I} := I(\overline{x})$ and $B_{\overline{I}} := [e_i, \; i \in \overline{I}]$. By Lemma 3, $\overline{x}$ and corresponding multipliers $\overline{\lambda} \in \mathbb{R}, 0 \leq \overline{\mu} \in \mathbb{R}^{\overline{I}}$ (by SC, $\overline{\mu} > 0$) are solutions of the KKT equations, (see (3))

$$M(A) \begin{pmatrix} x \\ -\lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \; , \quad \text{where} \quad M(A) = \begin{pmatrix} A & e & B_{\overline{I}} \\ e^T & 0 & 0 \\ B_{\overline{I}}^T & 0 & 0 \end{pmatrix} .$$

By LICQ and SOC the matrix $M(\overline{A})$ is nonsingular (see *e.g.*, [7, Ex. 12.20]). So, by continuity there is a neighborhood $\mathcal{N}_\varepsilon(\overline{A})$, $\varepsilon > 0$ such that for all $A \in \mathcal{N}_\varepsilon(\overline{A})$ the (rational) function

$$\begin{pmatrix} x(A) \\ -\lambda(A) \\ \mu(A) \end{pmatrix} = \begin{pmatrix} A & e & B_{\overline{I}} \\ e^T & 0 & 0 \\ B_{\overline{I}}^T & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

is well-defined and satisfies $\mu(A) > 0$ (recall $\mu(\overline{A}) = \overline{\mu} > 0$). Note, that by the condition $B_{\overline{I}}^T x = 0$ we have $I(x(A)) = \overline{I}$ and thus $R(x(A)) = S(x(A))$, *i.e.*, SC holds for $x(A)$. So, the solutions $x(A)$ are (locally unique) KKT points of $P_S(A)$. To show that $x(A)$ are ESS we have to show that also the second order condition SOC holds. This can be done by standart continuity arguments as in the proof of [7, Th.12.8].

□

In the case $R(\overline{x}) \subsetneqq S(\overline{x})$, at an ESS $\overline{x}$ of $\overline{A}$, the situation can be more complicated. In Example 1 we have seen that in this case, after a perturbation of $\overline{A}$ the ESS $\overline{x}$ may split into a whole set of (non-unique) local maximizers (i.e., the ESS can completely be lost). The next theorem however shows that locally the (unique) ESS behaves Lipschitz-stable if at the ESS $\overline{x}$ the stronger second order condition (SSOC) holds on the extended tangent space $T_{\overline{x}}$ (*cf.*, (7)):

$$d^T A d < 0 \ \ \forall 0 \neq d \in T_{\overline{x}}^+ \ . \tag{SSOC}$$

This follows by a result by Jittorntrum [10]. We again give the proof for our special case.

Theorem 3. *Let $\overline{x}$ be an ESS of $\overline{A} \in \mathcal{S}_m$ with $R(\overline{x}) \subsetneqq S(\overline{x})$ such that the condition SSOC holds. Then, there exist $\varepsilon, \delta > 0$ and a Lipschitz-function $x : \mathcal{N}_\varepsilon(\overline{A}) \to \mathcal{N}_\delta(\overline{x})$, $A \to x(A)$ with $x(\overline{A}) = \overline{x}$ and for any $A \in \mathcal{N}_\varepsilon(\overline{A})$ the vector $x(A)$ is an ESS of A and the unique local maximizer of A in $\mathcal{N}_\delta(\overline{x})$.*

Proof: By continuity any local maximizer $x = x(A) \approx \overline{x}$ must satisfy the KKT conditions with $R(\overline{x}) \subset R(x) \subset S(\overline{x})$ (we must have $[Ax]_i = \max_j [Ax]_j$ for all $i \in R(x)$ according to Lemma 2). So, in view of $I(x) = M \setminus R(x)$ the maximizer $x = x(A)$ must satisfy $M \setminus S(\overline{x}) \subset I(x) \subset I(\overline{x})$. Consequently, $x = x(A)$ must be a solution of one of the (finitely many) KKT systems:

$$F_I : \quad \begin{pmatrix} A & e & B_I \\ e^T & 0 & 0 \\ B_I^T & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ -\lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{where} \quad M \setminus S(\overline{x}) \subset I \subset I(\overline{x}) \ , \tag{11}$$

with corresponding multipliers $\lambda = \lambda(A) \in \mathbb{R}, \mu = \mu(A) \in \mathbb{R}_+^I$, and $B_I := [e_i, \ i \in I]$. By Lemma 5 the local solutions behave Lipschitz-continuous. So, we only have to show that under our assumptions, for any $A \in \mathcal{N}_\varepsilon(\overline{A})$ there exists a unique local maximizer $x(A)$.

Suppose to the contrary that there is a sequence $A_\nu \to A, \nu \to \infty$ such that A_ν has two different maximizers $x_\nu^1 \neq x_\nu^2$ near $\overline{x}$. By the Lipschitz continuity result in Lemma 5 we have $x_\nu^\rho \to \overline{x}, \ \rho = 1, 2$, for $\nu \to \infty$. By choosing appropriate subsequences wlog. we can assume that $I(x_\nu^1) =: I_1$ and $I(x_\nu^2) =: I_2$ holds with, $I_1 \neq I_2$ and $M \setminus S\overline{x}) \subset I_1, I_2 \subset I(\overline{x})$. So, the local maximizers $x_\nu^\rho, \ \rho = 1, 2$, are solutions of the corresponding KKT system

$$A_\nu x_\nu^\rho = \lambda_\nu^\rho - B_{I_\rho} \mu_\nu^\rho = 0, \quad e^T x_\nu^\rho = 1, \quad [x_\nu^\rho]_i = 0, \ i \in I_\rho, \ \rho = 1, 2 \ , \tag{12}$$

with $\mu_\nu^\rho \geq 0$. Since either $(x_\nu^1)^T A_\nu x_\nu^1 \leq (x_\nu^2)^T A_\nu x_\nu^2$ holds or the converse, by again choosing a subsequence we can assume

$$0 \leq (x_\nu^2)^T A_\nu x_\nu^2 - (x_\nu^1)^T A_\nu x_\nu^1 \quad \text{for all } \nu \ . \tag{13}$$

Now, let us define $d_\nu := \frac{x_\nu^2 - x_\nu^1}{\tau_\nu}$ with $\tau_\nu := \|x_\nu^2 - x_\nu^1\|$. Wlog. we can assume that the sequence d_ν converges, $d_\nu \to d, \ \|d\| = 1$. In view of $e^T d_\nu = 0$ and $[x_\nu^1]_i = 0, \ i \in I_1$ (see (12)) we find

$$[x_\nu^2]_i - [x_\nu^1]_i \geq 0 \quad \text{and thus} \quad [d_\nu]_i \geq 0 \ \ \forall i \in I_1 \quad \text{and also } [d_\nu]_i = 0, \ \ i \in M \setminus S(\overline{x}) \ , \tag{14}$$

in view of $M \setminus S(\overline{x}) \subset I_1, I_2$. By taking the limit $\nu \to \infty$ yields for d and its components $[d]_i$,

$$e^T d = 0, \quad [d]_i = 0, i \in M \setminus S(\overline{x}), \quad [d]_i \geq 0, \ i \in I_1 \ .$$

This implies $d \in T_{\overline{x}}^+$ (see (7)) . In view of (13), and using $-2(x_\nu^2 - x_\nu^1)^T B_{I_1} \mu_\nu^1 = -2 \sum_{i \in I_1} [x_\nu^2 - x_\nu^1]_i [\mu_\nu^1]_i \leq 0$ (by (14)) as well as the KKT conditions for x_ν^1, we obtain

$$\begin{array}{rcl} 0 & \leq & (x_\nu^2)^T A_\nu x_\nu^2 - (x_\nu^1)^T A_\nu x_\nu^1 \\ & = & 2(x_\nu^2 - x_\nu^1)^T A_\nu x_\nu^1 \ + \ (x_\nu^2 - x_\nu^1)^T A_\nu (x_\nu^2 - x_\nu^1) \\ & = & -2(x_\nu^2 - x_\nu^1)^T B_{I_1} \mu_\nu^1 \ + \ (x_\nu^2 - x_\nu^1)^T A_\nu (x_\nu^2 - x_\nu^1) \\ & \leq & (x_\nu^2 - x_\nu^1)^T A_\nu (x_\nu^2 - x_\nu^1) \end{array}$$

By dividing these relations by $\tau_\nu^2 > 0$ and letting $\nu \to \infty$, it follows

$$0 \leq d^T A d \quad \text{with} \quad d \in T_{\overline{x}}^+, \ d \neq 0 \ ,$$

contradicting the condition SSOC. □

Note that the only difference with the result in Theorem 2 is that in Theorem 3, the function $x(A)$ (possibly) is only Lipschitz continuous. We also provide an example.

Example 2. *[no SC, but second order condition on $T_{\overline{x}}^+$] The matrix* $\overline{A} := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ *has an ESS $\overline{x} = (1,0,0)$ satisfying $R(\overline{x}) = \{1\}$ and $S(\overline{x}) = \{1,2,3\}$. For this example we find (see (7)) $T_{\overline{x}}^+ = \{d \in \mathbb{R}^m \mid e^T d = d_1 + d_2 + d_3 = 0\}$ and then for any $d \in T_{\overline{x}}^+$, $d \neq 0$, in view of $d_1 = -d_2 - d_3$,*

$$d^T A d = d^T (0, d_1 + d_3, d_1 + d_2)^T = -d_2^2 - d_3^2 < 0 \ .$$

So, SSOC is satisfied and by the preceding theorem, locally, the ESS $\overline{x}$ behaves Lipschitz-stable after small perturbations of $\overline{A}$.

6 Genericity results for local maximizers

In optimization it is well-known that generically ("for a generic subset of problem instances") any local maximizer $\overline{x}$ is a nondegenerate strict local maximizer, *i.e.*, LICQ holds and the KKT condition is fulfilled with SC and SOC (see [11, Theorem 7.1.5]). We refer the reader to the landmark book [11] for genericity results in general nonlinear optimization.

We will formulate the genericity results specialized to our problem $P_S(A)$ and provide an easy and independent proof of such a genericity statement. This proof only makes use of the following basis result in differential geometry.

Lemma 6. *Let $p : \mathbb{R}^K \to \mathbb{R}$ be a polynomial mapping, $p \neq 0$. Then, the set of zeros of p, $p^{-1}(0) = \{x \in \mathbb{R}^K \mid p(x) = 0\}$, has (Lebesgue) measure zero in $\mathbb{R}^K$.*

Next we define what is meant by genericity. Note, that the set of problems $P_S(A),\ A \in \mathcal{S}_m$ can be identified with the set $\mathcal{Q} := \mathcal{S}_m$.

Definition 1. *We say that a property is generic in the problem set $\mathcal{S}_m$, if the property holds for a (generic) subset $\mathcal{Q}_r$ of $\mathcal{S}_m$ such that $\mathcal{Q}_r$ is open and $\mathcal{S}_m \setminus \mathcal{Q}_r$ has (Lebesgue) measure zero. (So, genericity implies density and stability of the set $\mathcal{Q}_r$ of "nice" problem instances.)*

The next theorem states that generically any local maximizer $\overline{x}$ of $P_S(A)$ is a nondegenerate (strict) local maximizer, *i.e.* an ESS with $R(\overline{x}) = S(\overline{x})$.

Theorem 4. *There is a generic subset $\mathcal{Q}_r \subset \mathcal{S}_m$ such that for any $A \in \mathcal{Q}_r$ the following holds: For any local maximizer $\overline{x}$ of $P_S(A)$ we have,*

$$(1)\quad R(\overline{x}) = S(\overline{x}), \quad \textit{i.e., SC is fulfilled and} \qquad (2)\quad \textit{SOC is satisfied.}$$

So, for any $A \in \mathcal{Q}_r$ any local maximizer $\overline{x}$ of $P_S(A)$ is an ESS point with $R(\overline{x}) = S(\overline{x})$.

Proof: (1): For a local maximizer $\overline{x}$ of $P_S(A)$, by Lemma 2, the condition $R(\overline{x}) \subset S(\overline{x})$ must be valid. Suppose now that this inclusion is strict i.e., $R(\overline{x}) \neq S(\overline{x})$. Then there exists some $j \in S(\overline{x}) \setminus R(\overline{x})$. This means that with $R := R(\overline{x})$ the point $0 < \overline{x}_R \in \mathbb{R}^{|R|}$ solves the system of linear equations

$$\begin{pmatrix} A_R \\ a_{j,R} \end{pmatrix} \overline{x} = m \begin{pmatrix} e_R \\ 1 \end{pmatrix} \qquad \text{with } m := \max_j [A\overline{x}]_j \ , \tag{15}$$

where $a_{j,R} := (a_{jl},\ l \in R)$. This implies that the determinant of the $(|R|+1)\times(|R|+1)$-matrix $\begin{pmatrix} A_R & e_R \\ a_{j,R} & 1 \end{pmatrix}$ is zero.

Consider now the polynomial function $p(A_R, a_{j,R}) := \det \begin{pmatrix} A_R & e_R \\ a_{j,R} & 1 \end{pmatrix}$. Since $p(I_R, 0) = 1$ this polynomial is nonzero and according to Lemma 6 for almost all $(A_R, a_{j,R}) \in \mathbb{R}^{|R|\cdot(|R|+1)}$ the relation $p(A_R, a_{j,R}) \neq 0$ holds, i.e., there is no solution of the equations (15). Moreover since the function $p(A_R, a_{j,R})$ is continuous, the set of parameters $(A_R, a_{j,R})$ with $p(A_R, a_{j,R}) \neq 0$ is open.

Since there is only a finite selection of subsets $R \subset M$ and elements $j \in M \setminus R$ possible, also the set of parameters A such that for all $R, j,\ R \subset M,\ j \in M \setminus R$, the condition $p(A_R, a_{j,R}) \neq 0$ holds, is generic. So, by construction, the condition $R(\overline{x}) \subsetneq S(\overline{x})$ is generically excluded.

(2): Now suppose that for a local maximizer $\overline{x}$ of $P_S(A)$ (by the above analysis we can assume $R(\overline{x}) = S(\overline{x})$) the condition SOC is not fulfilled, *i.e.*, $\overline{x}$ is not a strict local maximizer. In view of Lemma 4

$$\det(A_{R(\overline{x})}) = 0 \tag{16}$$

must be true. But, by defining the non-zero polynomial $p(A) := \det(A_{R(\overline{x})})$ and using Lemma 6 the condition (16) is excluded for almost all A. By noticing that also the condition $\det(A_{R(\overline{x})}) \neq 0$ is stable wrt. small perturbations of A the condition (16) is generically excluded.

□

A similar result is valid for the problem $P_B(A)$.

Theorem 5. *There is a generic subset $\mathcal{R}_r \subset \mathcal{S}_m$ such that for any $A \in \mathcal{R}_r$ all eigenvalues of A are simple. In particular, generically, the problem $P_B(A)$ has a unique solution.*

Proof: The proof follows from a more general stratification result for matrices (*cf.*, [1]). □

We conclude the paper with an observation. In Section 3 we have presented a matrix $\overline{A} \in \mathcal{S}_m$ with $(3^{\frac{1}{3}})^m$ strict local maximizers (exponential growth). Any of these ESS points $\overline{x}$ satisfies $R(\overline{x}) = S(\overline{x})$. We now might expect that for a generic set of $A \in \mathcal{S}_m$ (see Def. 1), such a large number of ESS is excluded. However this is not the case. By our stability result in Theorem 2 all these $(3^{\frac{1}{3}})^m$ ESS are locally stable, *i.e.*, (for fixed m) with some $\varepsilon > 0$, any matrix $A \in \mathcal{N}_\varepsilon(\overline{A})$ has $(3^{\frac{1}{3}})^m$ ESS.

References

[1] Arnold, V.I., *Singularity Theory*, Lond. Math. Society Lecture Notes Series 53, Cambridge University Press, (1981).

[2] Bomze I.M., *Non-cooperative two-person games in biology: A classification*, Int. J. of Game Theory, 15, Nr. 1, 31-57, (1986).

[3] Bomze I.M., *Evolution towards the maximum clique*, J. of Global Optimization, 10, 143-164, (1997).

[4] Bonnans J.F. and Shapiro A., *Perturbation analysis of optimization problems*, Springer Series in Operations Research, Springer Verlag, (2000).

[5] Cannings C. and Vickers G.T., *Patterns of ESS's II*, J. Theor, Biology, 132, 409-420, (1988).

[6] Fiacco A.V., *Introduction to sensitivity and stability analysis in nonlinear programming*, Academic Press, New York, (1983).

[7] Faigle U., Kern W., Still G., *Algorithmic Principles of Mathematical Programming*, Kluwer, Dordrecht, (2002).

[8] Golub G.H. and Van Loan Ch., Matrix computation, North Oxford Academy, (1987).

[9] Greenlee W.M. and Schaeffer A., *Iterative methods for eigenvalues of symmetric matrices as fixed point theorems*, Unpublished note, December 6, 2007.

[10] Jittorntrum K., *Solution point differentiability without strict complementarity in nonlinear programming*, in :A.V. Fiacco (ed.), Sensitivity, Stability and Parametric Analysis, Math. Progr. Study 21, 127-138, (1984).

[11] Jongen H.Th., Jonker P., Twilt F., *Nonlinear Optimization in finite Dimensions*, Kluwer, Dordrecht, (2000).

[12] Mandel, S.P.H., and Scheuer, P.A.G, *An inequality in population genetics*, Heredity 13, 519-524, (1959).

[13] Murty, K.G. and Kabadi, Santosh N., *Some NP-complete problems in quadratic and nonlinear programming*. Math. Programming 39, no. 2, 117-129, (1987).

[14] Smith J. Maynard, *The theory of games and the evolution of animal conflicts*, J. Theor. Biology, 47, 209-221, (1974).