Distance Computation from an Ellipsoid to a Linear or a Quadric Surface in $\mathbb{R}^n$

Alexei Yu. Uteshev and Marina V. Yashina

Faculty of Applied Mathematics, St. Petersburg State University
Universitetskij pr. 35, Petrodvorets, St.Petersburg, Russia
Alexei.Uteshev@pobox.spbu.ru

Abstract. Given the equations of the surfaces, our goal is to construct a univariate polynomial one of the zeroes of which coincides with the square of the distance between these surfaces. To achieve this goal we employ the Elimination Theory methods.

1 Problem Statement

Find the distance d from the ellipsoid

$$X^T \mathbf{A}_1 X + 2B_1^T X - 1 = 0 \tag{1}$$

a) to linear surface given by the system of equations

$$C_1^T X = 0, \dots, C_k^T X = 0 \tag{2}$$

b) to quadric

$$X^T \mathbf{A}_2 X + 2B_2^T X - 1 = 0. \tag{3}$$

Here $X = [x_1, \dots, x_n]^T$ is the column of variables, $\{B_1, B_2, C_1, \dots, C_k\} \subset \mathbb{R}^n$ are the given columns, and $C_1, \dots, C_k$ $(k \le n)$ are assumed to be linearly independent, $\mathbf{A}_1$ and $\mathbf{A}_2$ are the given symmetric matrices, and $\mathbf{A}_1$ is sign-definite.

Such problem arises in Computational Geometry [1,2], for instance, in the pattern recognition problem where one has to estimate the closeness of the objects given in n-dimensional parametric space.

The stated problem, being a problem of constrained optimization:

$$\min (X-Y)^T(X-Y) \text{ subject to } \begin{cases} X \in (1), Y \in (2) \text{ in the case of } a), \\ X \in (1), Y \in (3) \text{ in the case of } b), \end{cases}$$

can be reduced, via the conventional application of Lagrange multipliers method, to a problem of solving a system of algebraic equations. Thus, for instance, in the case of b)

$$\begin{cases} z - (X-Y)^T(X-Y) = 0 \\ X - Y - \lambda_1(\mathbf{A}_1 X + B_1) = \mathbf{O},\ -X + Y - \lambda_2(\mathbf{A}_2 Y + B_2) = \mathbf{O} \\ X^T \mathbf{A}_1 X + 2B_1^T X = 1,\ Y^T \mathbf{A}_2 Y + 2B_2^T Y = 1. \end{cases} \tag{4}$$

V.G. Ganzha, E.W. Mayr, and E.V. Vorozhtsov (Eds.): CASC 2007, LNCS 4770, pp. 392–401, 2007.

The main objective of this paper is to eliminate all the variables from this system except for z, i.e., to construct an algebraic equation $\mathcal{F}(z) = 0$ one of the zeros of which coincides with the square of the distance [3]. On evaluation of the latter, one can generically express the coordinates of the nearest points on the given surfaces as rational functions of this value.

2 Elimination Theory

The constructive realization of the declared procedure can be performed either via the Gröbner basis construction or with the aid of the classical Elimination Theory toolkit. From the latter the most suitable tool for solving our problem turns out to be the **discriminant**. For the (uni- or multivariate) polynomial $g(X) \in \mathbb{R}[X]$ its discriminant is formally and up to a multiple defined as

$$\mathcal{D}_X(g) \stackrel{def}{=} \prod_{j=1}^{N} g(\Lambda_j),$$

where $\{\Lambda_1, \ldots, \Lambda_N\}$ is a set of zeros (counted in accordance with their multiplicities) of the system

$$\frac{\partial g}{\partial x_1} = 0, \ldots, \frac{\partial g}{\partial x_n} = 0.$$

Discriminant can be expressed as a rational function of the coefficients of $g(X)$ with the aid of several determinantal representations. For instance, by the Bézout method [4,5]

$$\mathcal{D}_X(g) = \det \left[b_{\ell j} \right]_{\ell,j=0}^{N-1} . \qquad (5)$$

Here, for the univariate case and for $\deg g(X) = N + 1$, the element $b_{\ell j}$ stands for the coefficient of the remainder obtained on dividing $X^\ell g(X)$ by $g'(X)$:

$$X^\ell g(X) \equiv b_{\ell 0} + b_{\ell 1} X + \ldots + b_{\ell, N-1} X^{N-1} + q_\ell(X) g'(X), \ell \in \{0, \ldots, N-1\}.$$

As for the bivariate case, the element $b_{\ell j}$ of the matrix (5) is the coefficient of the **reduction** of the polynomial $\mathcal{M}_\ell(X) g(X)$ **modulo** $\partial g / \partial x_1$ **and** $\partial g / \partial x_2$:

$$\begin{aligned} \mathcal{M}_\ell(X) g(X) \equiv b_{\ell 0} \mathcal{M}_0(X) + \ldots + b_{\ell, N-1} \mathcal{M}_{N-1}(X) + \\ + q_{\ell 1}(X) \partial g / \partial x_1 + q_{\ell 2}(X) \partial g / \partial x_2 . \end{aligned}$$

Here $\{q_{\ell 1}(X), q_{\ell 2}(X)\} \subset \mathbb{R}[X]$, while $\{\mathcal{M}_\ell(X)\}_{\ell=0}^{N-1} \subset \mathbb{R}[X]$ is a set of the appropriately chosen power products in X. For the particular case of the polynomial standing as an argument for the discriminant function in Theorem 3, one should take $N = (n+1)^2$ and

$$\{\mathcal{M}_\ell(X)\}_{\ell=0}^{N-1} = \left\{ x_1^{j_1} x_2^{j_2} \,\middle|\, 0 \le j_1 < n+1, 0 \le j_2 \le 2(n - j_1) \right\} . \qquad (6)$$

The constructive reduction algorithm with respect to such a set was presented in [5].

If $\mathcal{D}_X(g) = 0$ then $g(X)$ possesses a multiple zero; if the latter is unique then it can be expressed rationally via the coefficients of the polynomial. This can be constructively performed with the aid of the minors of the determinant (5). For the univariate case this zero is given by

$$X = B_{N2}/B_{N1} \tag{7}$$

where B_{Nj} are the cofactors of the elements of the last row of the determinant (5). As for the bivariate polynomial from Theorem 3, let us reorder the power products of the set (6) in such a manner that $\mathcal{M}_0 = 1,\ \mathcal{M}_1 = x_1,\ \mathcal{M}_2 = x_2$ and denote by B_{Nj} the cofactors of the elements of the last row of the corresponding determinant (5). Then the components of the multiple zero can be expressed as

$$x_1 = B_{N2}/B_{N1},\ x_2 = B_{N3}/B_{N1}. \tag{8}$$

3 Distance to a Linear Surface

Theorem 1. *Construct the matrices* $\mathbf{C} \stackrel{def}{=} [C_1, \dots, C_k]$ *and* $\mathbf{G} \stackrel{def}{=} \mathbf{C}^T\mathbf{C}$ *(i.e.* $\mathbf{G}$ *is the Gram matrix for the columns* $C_1, \dots, C_k$*). The condition*

$$0 \le \left| \begin{array}{ccc} \mathbf{A}_1 & B_1 & \mathbf{C} \\ B_1^T & -1 & \mathbf{O} \\ \mathbf{C}^T & \mathbf{O} & \mathbf{O} \end{array} \right| \times \begin{cases} (-1)^{k-1} \ \mathit{if}\ \mathbf{A}_1\ \textit{is positive definite} \\ (-1)^n \ \mathit{if}\ \mathbf{A}_1\ \textit{is negative definite} \end{cases} \tag{9}$$

is the necessary and sufficient one for the linear surface (2) to intersect the ellipsoid (1); in this case one has $d = 0$*. If this intersection condition does not satisfied then the value* d^2 *coincides with the minimal positive zero of the equation*

$$\mathcal{F}(z) \stackrel{def}{=} \mathcal{D}_\mu \left(\mu^k \left| \begin{array}{ccc} \mathbf{A}_1 & B_1 & \mathbf{C} \\ B_1^T & -1 + \mu z & \mathbf{O} \\ \mathbf{C}^T & \mathbf{O} & \frac{1}{\mu}\mathbf{G} \end{array} \right| \right) = 0 \tag{10}$$

provided that this zero is not a multiple one.

Proof. **I. Finding the intersection condition.** Let us find first the critical value of[1] $V(X) = X^T\mathbf{A}X + 2B^TX - 1$ in the surface $\mathbf{C}^TX = \mathbf{O}$. The critical point of the Lagrange function

$$L = X^T\mathbf{A}X + 2B^TX - 1 - \nu_1 C_1^T X - \dots - \nu_k C_k^T X$$

satisfies the system of equations

$$2\mathbf{A}X + 2B - \mathbf{C}\left[\nu_1, \dots, \nu_k\right]^T = \mathbf{O},\ \mathbf{C}^TX = \mathbf{O}.$$

Therefrom

$$X = -\mathbf{A}^{-1}B + \frac{1}{2}\mathbf{A}^{-1}\mathbf{C}\left[\nu_1, \dots, \nu_k\right]^T \tag{11}$$

[1] To simplify the notation we will type matrices $\mathbf{A}$ and B without their subscript.

with

$$[\nu_1, \ldots, \nu_k]^T = 2\left(\mathbf{C}^T\mathbf{A}^{-1}\mathbf{C}\right)^{-1}\mathbf{C}^T\mathbf{A}^{-1}B. \quad (12)$$

Substitution of (12) into (11) yields

$$X_e = -\mathbf{A}^{-1}B + \mathbf{A}^{-1}\mathbf{C}\left(\mathbf{C}^T\mathbf{A}^{-1}\mathbf{C}\right)^{-1}\mathbf{C}^T\mathbf{A}^{-1}B$$

and the corresponding critical value of $V(X)$ subject to $\mathbf{C}^T X = \mathbf{O}$ equals

$$V(X_e) = -(B^T\mathbf{A}^{-1}B + 1 - B^T\mathbf{A}^{-1}\mathbf{C}(\mathbf{C}^T\mathbf{A}^{-1}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{A}^{-1}B).$$

With the aid of the Schur complement formula [6]:

$$\det\begin{pmatrix}\mathbf{U} & \mathbf{V}\\ \mathbf{S} & \mathbf{T}\end{pmatrix} = \det\mathbf{U}\det\left(\mathbf{T} - \mathbf{S}\mathbf{U}^{-1}\mathbf{V}\right) \quad (13)$$

(here $\mathbf{U}$ and $\mathbf{T}$ are square matrices and $\mathbf{U}$ is non-singular) one can transform the last expression into

$$V(X_e) = \frac{-\begin{vmatrix}\mathbf{C}^T\mathbf{A}^{-1}\mathbf{C} & \mathbf{C}^T\mathbf{A}^{-1}B\\ B^T\mathbf{A}^{-1}\mathbf{C} & B^T\mathbf{A}^{-1}B+1\end{vmatrix}}{\det(\mathbf{C}^T\mathbf{A}^{-1}\mathbf{C})} = \frac{(-1)^k\begin{vmatrix}\mathbf{A} & B & \mathbf{C}\\ B^T & -1 & \mathbf{O}\\ \mathbf{C}^T & \mathbf{O} & \mathbf{O}\end{vmatrix}}{\det(\mathbf{A})\det(\mathbf{C}^T\mathbf{A}^{-1}\mathbf{C})}. \quad (14)$$

If $V(X_e) = 0$ then the linear surface (2) is tangent to the ellipsoid (1) at $X = X_e$. Otherwise let us compare the sign of $V(X_e)$ with the sign of $V(X)$ at infinity. These signs will be distinct iff the considered surfaces intersect. If $\mathbf{A}$ is positive definite then $V_\infty > 0$, $\det(\mathbf{A}) > 0$ and $\det(\mathbf{C}^T\mathbf{A}^{-1}\mathbf{C}) > 0$. Therefore, $V(X_e) < 0$ iff the numerator in (14) is negative. This confirms (9). The case of negative definite matrix $\mathbf{A}$ is treated similarly.

II. Distance evaluation. Using the Lagrange multipliers method we reduce the constrained optimization problem to the following system of algebraic equations

$$X - Y - \lambda\mathbf{A}X - \lambda B = \mathbf{O} \quad (15)$$

$$X - Y + \frac{1}{2}\mathbf{C}[\lambda_1, \ldots, \lambda_k]^T = \mathbf{O} \quad (16)$$

$$X^T\mathbf{A}X + 2B^TX - 1 = 0 \quad (17)$$

$$\mathbf{C}^TY = \mathbf{O}. \quad (18)$$

We introduce also a new variable responsible for the critical values of the distance function:

$$z - (X-Y)^T(X-Y) = 0. \quad (19)$$

Our aim is to eliminate all the variables from the system (15)–(19) except for z. We express first X and Y from (15) and (16) (hereinafter $\mathbf{I}$ stands for the identity matrix of an appropriate order):

$$X = -\mathbf{A}^{-1}B - \frac{1}{2\lambda}\mathbf{A}^{-1}\mathbf{C}[\lambda_1, \ldots, \lambda_k]^T \quad (20)$$

$$Y = -\mathbf{A}^{-1}B - \frac{1}{2\lambda}(\mathbf{A}^{-1} - \lambda\mathbf{I})\mathbf{C}[\lambda_1, \ldots, \lambda_k]^T. \quad (21)$$

Then we substitute (21) into (18) with the aim to express $\lambda_1, \dots, \lambda_k$ via λ. This can be performed with the aid of the following matrix

$$\mathbf{M} \stackrel{def}{=} \frac{1}{\lambda}\mathbf{C}^T\mathbf{A}^{-1}\mathbf{C} - \mathbf{C}^T\mathbf{C} = \mu\mathbf{C}^T\mathbf{A}^{-1}\mathbf{C} - \mathbf{G}, \tag{22}$$

where $\mathbf{G}$ is the Gram matrix of the columns $C_1, \dots, C_k$ and $\mu \stackrel{def}{=} 1/\lambda$. Indeed, one has

$$\mathbf{M}[\lambda_1, \dots, \lambda_k]^T = -2\mathbf{C}^T\mathbf{A}^{-1}B \tag{23}$$

and, provided that $\mathbf{M}$ is non-singular,

$$[\lambda_1, \dots, \lambda_k]^T = -2\mathbf{M}^{-1}\mathbf{C}^T\mathbf{A}^{-1}B. \tag{24}$$

Now substitute (24) into (16) and then the obtained result into (19):

$$z - B^T\mathbf{A}^{-1}\mathbf{C}\mathbf{M}^{-1}\mathbf{G}\mathbf{M}^{-1}\mathbf{C}^T\mathbf{A}^{-1}B = 0. \tag{25}$$

Equation (25) is a rational one with respect to the variables μ and z.

To find an extra equation for these variables, let us transform (17) using (20) and (24)

$$\begin{aligned} 0 = X^T\mathbf{A}X + 2B^TX - 1 = -B^T\mathbf{A}^{-1}B - 1 + \\ + \mu B^T\mathbf{A}^{-1}\mathbf{C}\mathbf{M}^{-1}(\mu\mathbf{C}^T\mathbf{A}^{-1}\mathbf{C} - \mathbf{G} + \mathbf{G})\mathbf{M}^{-1}\mathbf{C}^T\mathbf{A}^{-1}B. \end{aligned}$$

Using (22) and (25), the last equation takes the form

$$\Psi(\mu, z) \stackrel{def}{=} -1 + \mu z - B^T\mathbf{A}^{-1}B + \mu B^T\mathbf{A}^{-1}\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^T\mathbf{A}^{-1}B = 0. \tag{26}$$

Therefore, the system (15)–(19) is reduced to (25)–(26). It can be verified that the left-hand side of (25) is just the derivative of that of (26) with respect to μ and, thus, it remains to eliminate μ from the system

$$\Psi(\mu, z) = 0, \ \Psi'_\mu(\mu, z) = 0.$$

This can be done with the help of discriminant – and that is the reason of its appearence in the statement of the theorem.

The Schur complement formula (13) helps once again in representing $\Psi(\mu, z)$ in the determinantal form:

$$\Psi(\mu, z) \equiv \frac{\left|\begin{array}{ccc} \mathbf{A} & B & \mathbf{C} \\ B^T & -1 + \mu z & \mathbf{O} \\ \mathbf{C}^T & \mathbf{O} & \frac{1}{\mu}\mathbf{G} \end{array}\right|}{\left|\begin{array}{cc} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^T & \frac{1}{\mu}\mathbf{G} \end{array}\right|} = \frac{\mu^k \left|\begin{array}{ccc} \mathbf{A} & B & \mathbf{C} \\ B^T & -1 + \mu z & \mathbf{O} \\ \mathbf{C}^T & \mathbf{O} & \frac{1}{\mu}\mathbf{G} \end{array}\right|}{\det(\mathbf{A})\det(\mathbf{M})}. \tag{27}$$

III. Finding the nearest points on the surfaces. Once the real zero $z = z_\star$ of (10) is evaluated, one can reverse the elimination scheme from part II of the proof in order to find the corresponding points $X_\star$ and $Y_\star$ on the surfaces.

For $z = z_\star$, the polynomial in μ standing in the numerator of (27) has a multiple zero $\mu = \mu_\star$. Provided that the multiple zero is unique, it can be expressed rationally in terms of the coefficients of this polynomial (and in $z_\star$) with the aid of (7). We substitute this value into (22) then resolve the linear system (23) with respect to $\lambda_1, \dots, \lambda_k$ and, finally, substitute the obtained numbers into (20) and (21).

However, this algorithm fails if for $\mu = \mu_\star$ the matrix $\mathbf{M}$ becomes singular. For explanation of the geometrical reason, one may recall that the distance between the surfaces may be attained not in a unique pair of points.

We avoid this case by imposing the simplicity restriction for the minimal zero of $\mathcal{F}(z)$ in the statement of the theorem. As a matter of fact, we are referring here to the empirical hypothesis that the conditions

$$\mathcal{D}_{x_1}(\mathcal{D}_{x_2}(g(x_1,x_2))) \ne 0 \text{ and } \mathcal{D}_{x_2}(\mathcal{D}_{x_1}(g(x_1,x_2))) \ne 0$$

are equivalent for the generic polynomial $g(x_1, x_2)$. For our particular case, the derivative of the determinant in the numerator of (27) with respect to z coincides with the denominator. □

Corollary 1. *If the system of columns $C_1, \dots, C_k$ is an orthonormal one then, by transforming the determinant in (10), one can diminish its order: the expression under discriminant can be reduced into*

$$\begin{vmatrix} \mathbf{A}_1 - \mu \mathbf{C}\mathbf{C}^T & B_1 \\ B_1^T & -1+\mu z \end{vmatrix} . \tag{28}$$

Example 1. Find the distance to the x_1-axis from the ellipsoid

$$7\,x_1^2 + 6\,x_2^2 + 5\,x_3^2 - 4\,x_1x_2 - 4\,x_2x_3 - 37\,x_1 - 12\,x_2 + 3\,x_3 + 54 = 0.$$

Solution. One can choose here $C_1 = [0,1,0]^T, C_2 = [0,0,1]^T$, then the determinant (28) takes the form

$$\begin{vmatrix} -7/54 & 1/27 & 0 & 37/108 \\ 1/27 & -1/9-\mu & 1/27 & 1/9 \\ 0 & 1/27 & -5/54-\mu & -1/36 \\ 37/108 & 1/9 & -1/36 & -1+\mu z \end{vmatrix} .$$

Equation (10)

$$\begin{aligned} \mathcal{F}(z) = {} & 516019098077413632\,z^4 - 15034745857812486912\,z^3 + \\ & + 95300876926947983328\,z^2 - 42103678084608945585 6\,z + \\ & + 237447832908365535785 = 0 \end{aligned}$$

has two real zeros: $z_1 = 0.05712805$ and $z_2 = 22.54560673$. Hence, the distance equals $\sqrt{z_1} \approx 0.23901475$.

Corollary 2. *The square of the distance from the origin* $X = \mathbf{O}$ *to the ellipsoid (1) coincides with the minimal positive zero of the equation*

$$\mathcal{F}(z) \stackrel{def}{=} \mathcal{D}_\mu \left(f(\mu)(\mu z - 1) - B_1^T q(\mathbf{A}_1, \mu) B_1 \right) = 0 \qquad (29)$$

provided that this zero is not a multiple one. Here $f(\mu) \stackrel{def}{=} \det(\mathbf{A}_1 - \mu \mathbf{I})$ *is the characteristic polynomial of the matrix* $\mathbf{A}_1$ *whereas* $q(\mathbf{A}_1, \mu)$ *stands for the adjoint matrix to the matrix* $\mathbf{A}_1 - \mu \mathbf{I}$.

Remark 1. For large n, one can compute $f(\mu)$ and $q(\mathbf{A}_1, \mu)$ simultaneously with the aid of the Leverrier–Faddeev method [7].

Remark 2. For the case $B_1 = \mathbf{O}$, one gets $\mathcal{F}(z) \equiv \mathcal{D}(f) \left[z^n f(1/z) \right]^2$. This corresponds to the well-known result that the distance to the ellipsoid $X^T \mathbf{A}_1 X = 1$ from its center coincides with the square root of the reciprocal of the largest eigenvalue of the matrix $\mathbf{A}_1$.

We exploit the result of the last corollary to elucidate the importance of the simplicity restriction imposed on the minimal positive zero for $\mathcal{F}(z)$; this assumption will also appear in the foregoing results.

Example 2. Find the distance from the origin to the ellipse

$$5/4\, x_1^2 + 5/4\, x_2^2 - 3/2\, x_1 x_2 - \alpha\, x_1 - \alpha\, x_2 + \alpha^2 - 1 = 0.$$

Here $\alpha > 0$ stands for parameter.

Solution. One can see that the given ellipse is obtained from the one centered at the origin by translation along its principal axis by the vector $[\alpha, \alpha]^T$. Let us investigate the dependence of the distance on α.

Polynomial (10)

$$\mathcal{F}(z) = \frac{1}{16} \frac{(z - 2\,\alpha^2 + 4\,\alpha - 2)\,(z - 2\,\alpha^2 - 4\,\alpha - 2)\,(6\,z + 4\,\alpha^2 - 3)^2}{(\alpha - 1)^6\,(\alpha + 1)^6}$$

possesses the zeros $z_1 = 2\,\alpha^2 + 4\,\alpha + 2,\ z_2 = 2\,\alpha^2 - 4\,\alpha + 2,\ z_3 = -2/3\,\alpha^2 + 1/2$, and, for any specialization of the parameter, the value d^2 will be among these values.

Furthermore, $z_3 = \min\{z_1, z_2, z_3\}$ for $\alpha\ \in]0, \sqrt{3}/2]$. Nevertheless, for $\alpha \in]3/4, \sqrt{3}/2]$ the square of the distance is calculated by the formula $d^2 = z_2$.

Explanation for this phenomenon is as follows: the multiple zero z_3 corresponds to the pair of points $[x_1, x_2]^T$ on ellipse. These points are real for $\alpha \le 3/4$ and imaginary (complex-conjugate) for $\alpha > 3/4$.

4 Distance to a Quadric

Consider first the case of surfaces centered at the origin: $B_1 = \mathbf{O},\ B_2 = \mathbf{O}$.

Theorem 2. *The surfaces* $X^T\mathbf{A}_1X = 1$ *and* $X^T\mathbf{A}_2X = 1$ *intersect iff the matrix* $\mathbf{A}_1 - \mathbf{A}_2$ *is not sign-definite. If this condition is not satisfied then the value* d^2 *coincides with the minimal positive zero of the equation*

$$\mathcal{F}(z) \stackrel{def}{=} \mathcal{D}_\lambda(\det(\lambda\mathbf{A}_1 + (z-\lambda)\mathbf{A}_2 - \lambda(z-\lambda)\mathbf{A}_1\mathbf{A}_2)) = 0 \tag{30}$$

provided that this zero is not a multiple one.

Remark 3. We failed to establish the authors of the intersection condition from the above theorem. However, this condition should be treated as "well-known" since it is contained as an exercise in the problem book [8]. The other assertion of Theorem 2 follows from

Theorem 3. *The surfaces (1) and (3) intersect iff among the real zeros of the equation*

$$\varPhi(z) \stackrel{def}{=} \mathcal{D}_\lambda\left(\det\left(\begin{bmatrix}\mathbf{A}_2 & B_2\\ B_2^T & -1-Z\end{bmatrix} - \lambda\begin{bmatrix}\mathbf{A}_1 & B_1\\ B_1^T & -1\end{bmatrix}\right)\right) = 0$$

there are the values of different signs or 0*. If this condition is not satisfied then the value* d^2 *coincides with the minimal positive zero of the equation*

$$\mathcal{F}(z) \stackrel{def}{=} \tag{31}$$

$$\stackrel{def}{=} \mathcal{D}_{\mu_1,\mu_2}\left(\det\left(\mu_1\begin{bmatrix}\mathbf{A}_1 & B_1\\ B_1^T & -1\end{bmatrix} + \mu_2\begin{bmatrix}\mathbf{A}_2 & B_2\\ B_2^T & -1\end{bmatrix} - \begin{bmatrix}\mathbf{A}_2\mathbf{A}_1 & \mathbf{A}_2B_1\\ B_2^T\mathbf{A}_1 & B_2^TB_1 - \mu_1\mu_2 z\end{bmatrix}\right)\right) = 0$$

provided that this zero is not a multiple one.

Proof. is sketched as it is similar to that of Theorem 1. Intersection condition is a result of the following considerations. Extrema of the function $X^T\mathbf{A}_2X + 2B_2^TX - 1$ on the ellipsoid (1) are all of the similar sign iff the surfaces (1) and (3) do not intersect. We state the problem of finding the extremal **values** of $X^T\mathbf{A}_2X + 2B_2^TX - 1$ subject to (1), then apply the Lagrange multipliers method and finally eliminate all the variables except for z from the obtained algebraic system coupled with the equation $X^T\mathbf{A}_2X + 2B_2^TX - 1 - z = 0$.

To prove the second part of the theorem denote

$$\mathbf{M} \stackrel{def}{=} \mathbf{I} - \frac{1}{\lambda_1}\mathbf{A}_1^{-1} - \frac{1}{\lambda_2}\mathbf{A}_2^{-1},\ Q \stackrel{def}{=} -\mathbf{A}_1^{-1}B_1 + \mathbf{A}_2^{-1}B_2$$

and transform the equations of the system (4) into

$$X = -\mathbf{A}_1^{-1}B_1 + \frac{1}{\lambda_1}\mathbf{A}_1^{-1}\mathbf{M}^{-1}Q,\ Y = -\mathbf{A}_2^{-1}B_2 - \frac{1}{\lambda_2}\mathbf{A}_2^{-1}\mathbf{M}^{-1}Q \tag{32}$$

$$-B_j^T\mathbf{A}_j^{-1}B_j + \frac{1}{\lambda_j^2}Q^T\mathbf{M}^{-1}\mathbf{A}_j^{-1}\mathbf{M}^{-1}Q - 1 = 0 \text{ for } j\in\{1,2\} \tag{33}$$

$$z - Q^T\mathbf{M}^{-2}Q = 0. \tag{34}$$

On multiplying equations (33) by λ_j and using (34), we get

$$-\lambda_1 B_1^T \mathbf{A}_1^{-1} B_1 - \lambda_2 B_2^T \mathbf{A}_2^{-1} B_2 - Q^T \mathbf{M}^{-1} Q - \lambda_1 - \lambda_2 + z = 0. \qquad (35)$$

It can be verified that the derivative of the left-hand side of (35) with respect to λ_j coincides with that one of (33). Substitution $\mu_1 = 1/\lambda_2,\ \mu_2 = 1/\lambda_1$ and the use of the Schur complement formula (13) enable one to reduce (35) to the determinantal representation from (31). □

Example 3. Find the distance between the ellipsoids

$$7\,x_1^2 + 6\,x_2^2 + 5\,x_3^2 - 4\,x_1x_2 - 4\,x_2x_3 - 37\,x_1 - 12\,x_2 + 3\,x_3 + 54 = 0$$

$$\text{and } 189\,x_1^2 + x_2^2 + 189\,x_3^2 + 2\,x_1x_3 - x_2x_3 - 27 = 0$$

and establish the coordinates of their nearest points.

Solution. Intersection condition from Theorem 3 is not satisfied: the sixth-order polynomial $\Phi(z)$ has all its real zeros positive. To compute the discriminant (31) we represent it as the determinant (5) of the order $N = 16$. The twenty-fourth-order polynomial $\mathcal{F}(z)$, with integer coefficients of the orders up to 10^{188}, has eight positive zeros $z_1 \approx 1.35377, \dots, z_8 \approx 111.74803$. Thus, the distance between the given ellipsoids equals $\sqrt{z_1} \approx 1.16351$.

For the obtained value of z_1, the polynomial in μ_1 and μ_2 from (31) possesses a multiple zero which can be expressed rationally in terms of z_1 with the aid of the minors of the determinant (5) by (8). Substitution of the obtained values $\lambda_1 \approx 5.75593,\ \lambda_2 \approx -0.45858$ into (32) yields the coordinates of the nearest points on the given ellipsoids:

$$X \approx [1.52039,\ 1.50986,\ 0.12623]^T,\ Y \approx [0.36100,\ 1.48490,\ 0.03152]^T.$$

Remark 4. It turns out that generically the degree of the polynomial $\mathcal{F}(z)$ is given by the following table

Formula	(10)	(29)	(30)	(31)
$\deg \mathcal{F}(z)$	$2k$	$2n$	$n(n+1)$	$2n(n+1)$

Formulas from the third and the fourth column are valid on excluding the extraneous factor from $\mathcal{F}(z)$ (in the case of the fourth column the mentioned factor is responsible for the equivalence of the transfer from the representation (35) to that one of (31)).

5 Conclusions

We have treated the problem of distance evaluation between algebraic surfaces in $\mathbb{R}^n$ via inversion of the traditional approach:

$$\text{nearest points} \ \rightarrow \ \text{distance}.$$

This has been performed via introduction of an extra variable responsible for the critical values of distance function and application of the Elimination Theory methods. It happens that the discriminant is fully responsible for everything: with its help it is not only possible to deduce a univariate polynomial equation for the square of the distance but also to express (Theorem 3) the necessary and sufficient condition for the intersection of the surfaces.

The proposed approach might be especially useful for the optimization problems connected with the parameter dependent surfaces, for instance, for finding an ellipsoid approximating a set of points in $\mathbb{R}^n$.

References

1. Schneider, P.J., Eberly, D.H.: Geometric Tools for Computer Graphics. Elsevier, San Francisco (2003)
2. Lin, A., Han, S.-P.: On the distance between two ellipsoids. SIAM J. on Optimization 13, 298–308 (2002)
3. Uteshev, A.Y., Cherkasov, T.M.: The search for maximum of a polynomial. J. Symbolic Computation 25, 587–618 (1998)
4. Kalinina, E.A., Uteshev, A.Y.: Elimination Theory (in Russian). St. Petersburg State University (2002)
5. Bikker, P., Uteshev, A.Y.: On the Bézout construction of the resultant. J. Symbolic Computation 28, 45–88 (1999)
6. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1986)
7. Faddeev, D.K., Faddeeva, V.N.: Computational Methods of Linear Algebra. Freeman, San Francisco (1963)
8. Proskuryakov, I.V.: Problems in Linear Algebra. Mir, Moscow (1978)