Elemente der Mathematik (Online first)
DOI 10.4171/EM/480

Published by EMS Press

Regression ellipses via hyperbolic geometry

Lilija Naiwert and Karlheinz Spindler

Lilija Naiwert received three years of training as a retail management assistant and then worked for three years in a clothing store before deciding to study mathematics. She enrolled in the Applied Mathematics line of study at the Hochschule RheinMain where she received her Bachelor's degree in 2017, her thesis being on parameter estimation problems in image processing and orbit determination. Two years later, she received her Master's degree, with a thesis in which a study of the critical points of multivariate polynomials was applied to a problem in stability analysis. Since graduating, she has worked as an actuary in a business consulting firm.

Karlheinz Spindler studied mathematics, mechanics and history at the Technische Hochschule Darmstadt. After obtaining his Ph.D. in mathematics, he spent two years as a visiting assistant professor at Louisiana State University in Baton Rouge (USA) and then worked for five years in the Flight Dynamics Department of the European Space Operations Centre (ESOC) in Darmstadt. He currently teaches mathematics and data processing at the Hochschule RheinMain (formerly Fachhochschule Wiesbaden). In addition to his research and teaching activities, he is also the founder and co-organizer of a continuing education seminar on technical, didactical and historical aspects of mathematics.

Dedicated to Prof. Karl Heinrich Hofmann on the occasion of his 90th birthday

We study the problem of finding a regression ellipse to fit a given family of data points (x_i, y_i), where $1 \le i \le n$, i.e., an ellipse which matches these data points "as well as possible". Since each ellipse can be uniquely specified by an equation of the form

$$ax^2 + bxy + cy^2 + dx + ey + f = 0, \quad \text{where } 4ac - b^2 = 1,$$

Die bestmögliche Schätzung von Systemparametern aus Messungen ist eine in vielen Anwendungssituationen auftretende Fragestellung. Erfüllen die zu schätzenden Parameter Nebenbedingungen, treten sie also als Elemente einer nichtlinearen Mannigfaltigkeit auf, so treffen bei dieser Fragestellung verschiedene Gebiete der Mathematik zusammen: Optimierung, Differentialgeometrie, Statistik. Der vorliegende Artikel zeigt am konkreten Beispiel der Bestimmung von Ausgleichsellipsen, wie sich differentialgeometrische Ideen zur eleganten Lösung eines Parameterschätzproblems nutzen lassen.

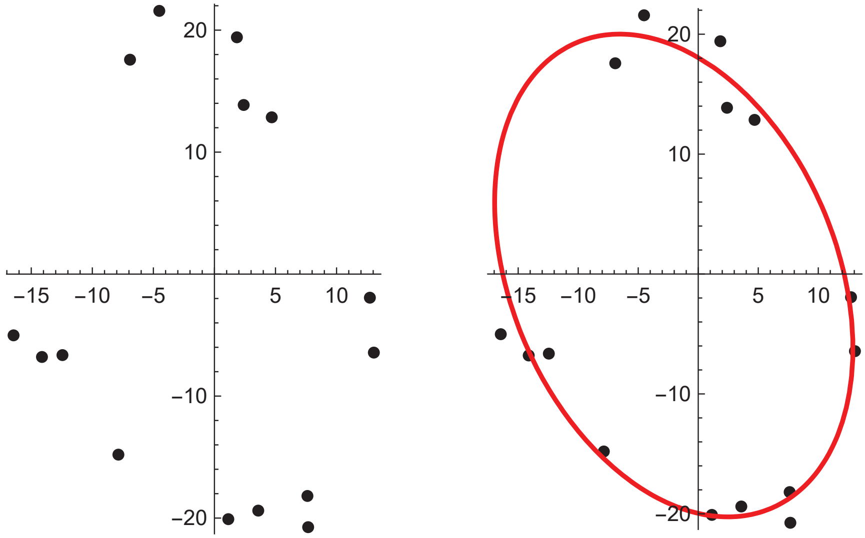

Figure 1. Family of data points (left) and regression ellipse through these points (right).

we can choose as a criterion for the goodness of fit the quantity

$$\sum_{i=1}^{n}(ax_i^2 + bx_i y_i + cy_i^2 + dx_i + ey_i + f)^2. \tag{1}$$

Thus we seek to minimize expression (1) amongst all arguments (a, b, c, d, e, f) in $\mathbb{R}^6$ subject to the constraint $4ac - b^2 = 1$. This problem can be solved using the method of Lagrange multipliers (see [17, Vol. 2, problem (98.37)]); an example for the solution is shown in Figure 1.

In this paper, we want to present an alternative solution which yields an elegant application of a rather general approach to regression problems on manifolds. To present this approach, let us summarize some key facts from regression analysis (cf. [5]). The simplest regression problem takes the form of an overdetermined system $Ax = b$ of linear equations, where $x \in \mathbb{R}^m$, $b \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times m}$ with $n > m$. We interpret x as a vector of m parameters and the n scalar equations constituting the system as measurements which are used to estimate the parameters. Since there are more measurements than parameters and since measurement errors are unavoidable, we can only hope to solve the system $Ax = b$ in an approximate way, namely in the sense of minimizing $\|Ax - b\|$ with respect to some norm on $\mathbb{R}^n$. Now if this norm is Euclidean and if we choose another Euclidean norm on $\mathbb{R}^m$, then a vector $x \in \mathbb{R}^m$ minimizes $\|Ax - b\|$ if and only if it satisfies the normal equation $A^T Ax = A^T b$, where the transpose of A is taken with respect to the inner products underlying the chosen norms. If A has rank m (which means that the measurements contain enough information to estimate the parameters), then x is uniquely given by

$$x = (A^T A)^{-1} A^T b. \tag{2}$$

Next we turn to nonlinear regression problems. Again, we seek a parameter vector $x \in \mathbb{R}^m$ which, ideally, satisfies an equation $f(x) = \mu$, where $\mu \in \mathbb{R}^n$ is a given vector of measurement values and where the components of the function $f: \mathbb{R}^m \to \mathbb{R}^n$ are interpreted as measurements. We reduce such a nonlinear regression problem to a sequence of linear regression problems as follows. Given an estimate $x^{(0)}$ for the parameter vector in question, we ask which update δx should be chosen such that the new estimate $x^{(0)} + \delta x$ satisfies the equation $f(x) = \mu$ as well as possible. If we assume that the original estimate is not too bad, i.e., not too far away from the optimal choice for x, we can linearize f about $x^{(0)}$ and thus have to satisfy the approximate equations $\mu \approx f(x^{(0)} + \delta x) \approx f(x^{(0)}) + f'(x^{(0)})\,\delta x$. This amounts to solving the linear regression problem

$$A\,\delta x = \rho, \quad \text{where } A := f'(x^{(0)}) \text{ and } \rho := \mu - f(x^{(0)}). \tag{3}$$

The vector ρ is called the residual vector; it is the difference between the vector μ of actually obtained measurements and the vector $f(x^{(0)})$ of theoretically expected measurements based on the assumption that $x^{(0)}$ is the true parameter vector. Applying (2) to (3) yields $\delta x = (A^T A)^{-1} A^T \rho$ and then $x^{(1)} := x^{(0)} + \delta x$ as an updated (and, as we hope, improved) estimate for the true parameter vector x. This procedure needs to be iterated, as in each step we slightly change the problem by linearizing about the currently best estimate, ignoring higher-order terms. In typical situations, this iterative estimation scheme converges fast, provided that the initial estimate is not too far off.

Finally, we turn to a regression problem of the form $f(p) = \mu$ in which the estimation parameter p is not a vector in $\mathbb{R}^m$, but an element of a manifold M. (In practical applications, M is typically an embedded submanifold of some Euclidean space so that we can treat p as a vector subject to certain constraints.) In order to adapt the iterative scheme described before, we assume that M is equipped with a Riemannian metric. We proceed as before, with the following modifications.

- Given an estimate p (which is constrained to lie in a manifold M), we only allow updates δp which are constrained to lie in the tangent space $T_p M$ of M at p, thereby solving a linear regression problem using the linearization $f'(p): T_p M \to \mathbb{R}^n$ of f at p and using the Euclidean norm on $T_p M$ induced by the Riemannian metric on M.
- Once the update $\delta p \in T_p M$ is found, it does not make sense to replace the old estimate p by the new estimate $p + \delta p$, not even if M is an embedded submanifold. (In this setting, $p + \delta p$ is a well-defined element of the ambient space, but not an element of M.) Thus the simple linear update step $x^{(0)} \mapsto x^{(0)} + \delta x$ encountered before is replaced by the nonlinear update step $p \mapsto \exp_p(\delta p)$, where $\exp_p$ is the exponential function of the manifold M at p, defined by $\exp_p(v) = \gamma(1)$, where γ is the unique geodesic on M satisfying $\gamma(0) = p$ and $\dot\gamma(0) = v$. Intuitively, the update vector δp is wrapped around M to find a new estimate; this is shown in Figure 2.

To apply the regression algorithm just described to the problem of regression ellipses, it will be convenient to write $b = \beta\sqrt{2}$ and to change the normalization constant to the value 2. Thus we seek the equation of the regression ellipse in the form

$$ax^2 + \sqrt{2}\beta xy + cy^2 + dx + ey + f = 0$$

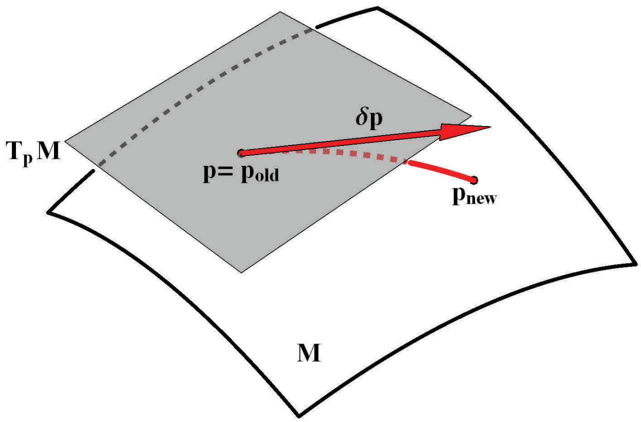

Figure 2. Nonlinear update step to improve a parameter estimate in an estimation problem on a Riemannian manifold.

subject to the constraint

$$1 = 2ac - \beta^2 = [a \;\; \beta \;\; c] \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ \beta \\ c \end{bmatrix}.$$

Rotating the coordinate system by 45° about the axis spanned by the vector $(1, 0, 1)^T$, i.e., introducing the new coordinates

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = T^{-1} \begin{bmatrix} a \\ \beta \\ c \end{bmatrix}, \quad \text{where } T := \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix},$$

the equation $2ac - \beta^2 = 1$ becomes $x^2 + y^2 - z^2 = -1$, which is the equation of a two-sheeted hyperboloid H. Thus we want to minimize the function

$$\sum_{i=1}^{n} \left(\frac{x+z}{\sqrt{2}} \cdot x_i^2 + \sqrt{2} y \cdot x_i y_i + \frac{z-x}{\sqrt{2}} \cdot y_i^2 + d x_i + e y_i + f \right)^2$$

subject to the constraint $x^2 + y^2 - z^2 = -1$. This is a minimization problem on the manifold $M := H \times \mathbb{R}^3 = \{(x, y, z, d, e, f) \in \mathbb{R}^6 \mid x^2 + y^2 - z^2 = -1\}$. In order to apply the method described before, we need to equip M with a Riemannian metric. To do so, we take the product metric of a Riemannian metric on H with the metric on $\mathbb{R}^3$ induced by the standard Euclidean structure. As a Riemannian metric on H, however, we do not use the metric induced by the canonical inner product on $\mathbb{R}^3$, but the one induced by the Minkowski bilinear form

$$\sigma\left(\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \right) := v_1 w_1 + v_2 w_2 - v_3 w_3,$$

which makes H into a Riemannian manifold of constant curvature -1 isometric to the hyperbolic plane. Let us describe the details. Writing $p = (x, y, z)$, we have

$$\begin{aligned} H &= \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 = -1\} \\ &= \{p \in \mathbb{R}^3 \mid \sigma(p, p) = -1\}. \end{aligned}$$

If $t \mapsto \xi(t) = (x(t), y(t), z(t))$ is a curve in H, then $\sigma(\xi(t), \xi(t)) \equiv -1$ and consequently $\sigma(\xi(t), \dot{\xi}(t)) \equiv 0$, which implies that the tangent space of H at a point $p = (x, y, z) \in H$ is given by

$$T_p H = \{v \in \mathbb{R}^3 \mid \sigma(p, v) = 0\}.$$

Now if $p = (x, y, z) \in H$ and $v = (v_1, v_2, v_3)^T \in T_p H$, then $x^2 + y^2 - z^2 = -1$ and $xv_1 + yv_2 - zv_3 = 0$, whence $z \neq 0$ and $v_3 = (xv_1 + yv_2)/z$. This implies that

$$\begin{aligned} \sigma(v, v) &= v_1^2 + v_2^2 - v_3^2 = v_1^2 + v_2^2 - \frac{(xv_1 + yv_2)^2}{z^2} \\ &= \frac{z^2 v_1^2 + z^2 v_2^2 - (xv_1 + yv_2)^2}{z^2} \\ &= \frac{(z^2 - x^2)v_1^2 + (z^2 - y^2)v_2^2 - 2xyv_1v_2}{z^2} \\ &= \frac{(y^2 + 1)v_1^2 + (x^2 + 1)v_2^2 - 2xyv_1v_2}{z^2} \\ &= \frac{v_1^2 + v_2^2 + (yv_1 - xv_2)^2}{z^2} \end{aligned}$$

and hence that $\sigma(v, v) > 0$ whenever $v \neq 0$. Thus σ is positive definite on each tangent space of H and therefore induces a Riemannian metric on H (which, as a matter of fact, makes H isometric to the hyperbolic plane; cf. [14, 15]). With this metric, the unique geodesic γ originating at a point p with a unit velocity vector v is given by

$$\gamma(t) := \cosh(t)p + \sinh(t)v. \tag{4}$$

Let us prove this claim! We show first that γ is indeed a curve within H. Noting that $\sigma(p, p) = -1$ because $p \in H$, that $\sigma(p, v) = 0$ because $v \in T_p H$ and that $\sigma(v, v) = 1$ because v has unit length, we find by bilinearity that $\sigma(\gamma(t), \gamma(t)) = \cosh(t)^2\sigma(p, p) + 2\sinh(t)\cosh(t)\sigma(p, v) + \sinh(t)^2\sigma(v, v) = -\cosh(t)^2 + \sinh(t)^2 = -1$, which implies that $\gamma(t) \in H$ for all t. Clearly, $\gamma(0) = p$ and $\dot{\gamma}(0) = v$. Finally, we have $\sigma(\gamma(t), v) = 0$ for all $v \in T_{\gamma(t)}H$, which implies that $\ddot{\gamma}(t) = \gamma(t) \in (T_{\gamma(t)}H)^{\perp}$ for all t and hence that the acceleration vector $\ddot{\gamma}$ has no tangent component, revealing γ to be a geodesic. The reason why we equip H not with the metric induced by $\mathbb{R}^3$, but with the hyperbolic metric, is exactly the existence of the explicit formula (4) for the geodesics on H. Since the parameters $(-x, -y, -z, -d, -e, -f)$ and (x, y, z, d, e, f) determine the same ellipse, we may assume that $z > 0$ and hence $z = \sqrt{1 + x^2 + y^2}$, which leaves only five estimation parameters (x, y, d, e, f). Thus the algorithm to determine the regression ellipse through points (x_i, y_i) (where $1 \leq i \leq n$) can be succinctly formulated as follows.

- Given an estimate $(x, y, z, d, e, f) \in M$, where $z = \sqrt{1 + x^2 + y^2}$, write

$$p := (x, y, z) \in H \quad \text{and} \quad q := (d, e, f) \in \mathbb{R}^3,$$

determine the residual vector $\rho \in \mathbb{R}^n$ whose i-th entry is

$$\rho_i := -\Big(\frac{x+z}{\sqrt{2}} \cdot x_i^2 + \sqrt{2}y \cdot x_i y_i + \frac{z-x}{\sqrt{2}} \cdot y_i^2 + dx_i + ey_i + f\Big)$$

and the matrix $A \in \mathbb{R}^{n \times 5}$ whose i-th row is

$$\Big(\frac{(x_i^2 - y_i^2) + (x_i^2 + y_i^2) \cdot x/z}{\sqrt{2}}, \sqrt{2} \cdot x_i y_i + \frac{(x_i^2 + y_i^2) \cdot y/z}{\sqrt{2}}, x_i, y_i, 1\Big)$$

and let $\Delta := (A^T A)^{-1} A^T \rho$. (Note that the i-th row of A consists of the partial derivatives of the i-th measurement with respect to x, y, d, e and f.)
- Write $\Delta = (\delta x, \delta y, \delta d, \delta e, \delta f)$ and let $\delta z := (x \cdot \delta x + y \cdot \delta y)/z$. Moreover, write $\delta p := (\delta x, \delta y, \delta z)^T$ and $\delta q := (\delta d, \delta e, \delta f)^T$ and perform the update steps

$$p_{\text{new}} := \cosh(\|\delta p\|) \cdot p + \sinh(\|\delta p\|) \cdot \frac{\delta p}{\|\delta p\|} \quad \text{and} \quad q_{\text{new}} := q + \delta q,$$

where the norm $\|\cdot\|$ is the one on $T_p H$ so that

$$\|\delta p\|^2 = \frac{(\delta x)^2 + (\delta y)^2 + (y\,\delta x - x\,\delta y)^2}{z^2}.$$

- Write $p_{\text{new}} = (x_{\text{new}}, y_{\text{new}}, z_{\text{new}})$ and $q_{\text{new}} = (d_{\text{new}}, e_{\text{new}}, f_{\text{new}})$ and replace the old estimate (x, y, z, d, e, f) by the new estimate $(x_{\text{new}}, y_{\text{new}}, z_{\text{new}}, d_{\text{new}}, e_{\text{new}}, f_{\text{new}})$.

Figure 3 shows an application of this algorithm. As an initial estimate $p^{(0)}$, we choose an ellipse passing through five of the given data points (shown in blue). The algorithm then yields successively the estimates $p^{(1)}$ (green), $p^{(2)}$ (orange) and $p^{(3)}$ (red). Note that $p^{(2)}$ and $p^{(3)}$ can hardly be distinguished, which shows that the algorithm converges fast to the regression ellipse through the given data points. The transition from an estimate $p^{(i)}$ to the next estimate $p^{(i+1)}$ takes place along a geodesic on M; this is shown on the right-hand side of Figure 3. Note that, instead of thinking of M as a manifold of sextuplets (x, y, z, d, e, f) *describing* ellipses, we may think of M as a manifold whose points *are* ellipses. This point of view is made tangible by viewing geodesics in M as time-varying families of ellipses as shown in Figure 3.

There is a rather extensive literature on the determination of regression ellipses (see, e.g., [1, 8–11, 19, 20]), and the optimization criterion used in the above example is not the only possible one. A more geometrically (rather than algebraically) motivated criterion for the goodness of fit is the sum of the squared distances from the given data points to the sought ellipse, and we want to show how our approach (which is concerned with modeling the parameter space rather than specifying the quantity to be minimized) can be used with this criterion as well. If a point Q on a given ellipse $aX^2 + bXY + cY^2 + dX + eY + f = 0$ has minimal distance to a given point P, then the vector $\overrightarrow{PQ}$ is perpen-

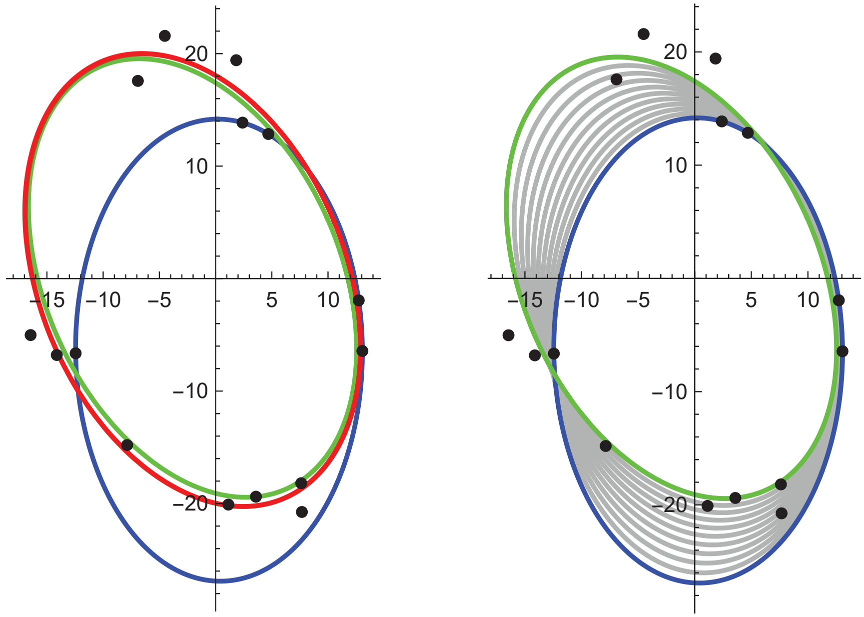

Figure 3. Iterative determination of the regression ellipse (left) and "points" (i.e., ellipses) on the geodesic joining the initial estimate and the resulting updated estimate (right).

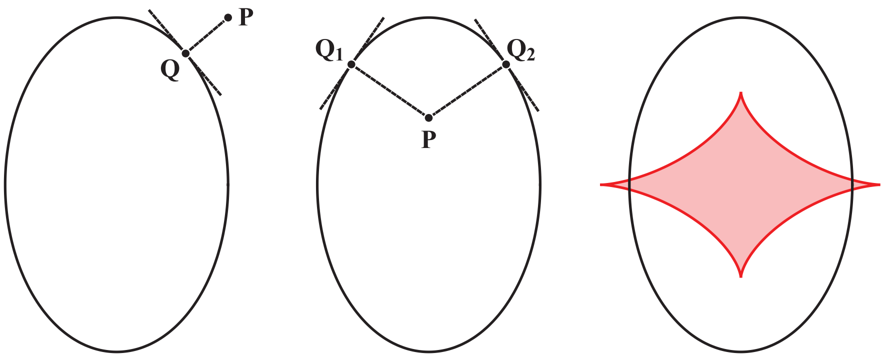

Figure 4. Left: Best approximation Q to a given point P on a given ellipse. Center: Point P with two best approximations Q_1 and Q_2. Right: area (red) of those points which admit four normal lines to the ellipse.

dicular to the tangent to the ellipse at Q; this can be easily seen by geometric reasoning or by applying the method of Lagrange multipliers. (See Figure 4. Note that such a best approximation is not necessarily uniquely determined. This is related to Apollonius' problem of determining the number of normal lines from a given point to a given ellipse; cf. [16, (98.9)].)

Applying this observation with any of the data points $P = (x_i, y_i)$ and writing $Q = (u_i, v_i)$ for the best approximation of P on the ellipse $aX^2 + bXY + cY^2 + dX +$

$eY + f = 0$, we find that

$$\begin{aligned} 0 &= (u_i - x_i)(bu_i + 2cv_i + e) - (v_i - y_i)(2au_i + bv_i + d) \\ &= bu_i^2 + (2c - 2a)u_i v_i - bv_i^2 + (-bx_i + 2ay_i + e)u_i \\ &\quad + (-2cx_i + by_i - d)v_i + (dy_i - ex_i). \end{aligned} \tag{5}$$

Moreover, since (u_i, v_i) is a point on the ellipse, we have

$$au_i^2 + bu_i v_i + cv_i^2 + du_i + ev_i + f = 0. \tag{6}$$

Rewriting equations (5) and (6) in terms of (x, y, z) instead of (a, b, c) results in the equations

$$0 = \frac{x + z}{\sqrt{2}} u_i^2 + \sqrt{2} y u_i v_i + \frac{z - x}{\sqrt{2}} v_i^2 + du_i + ev_i + f \tag{7}$$

and

$$\begin{aligned} 0 &= \sqrt{2} y u_i^2 - 2\sqrt{2} x u_i v_i - \sqrt{2} y v_i^2 \\ &\quad + \big(-\sqrt{2} y x_i + \sqrt{2}(x + z) y_i + e\big) u_i \\ &\quad + \big(-\sqrt{2}(z - x) x_i + \sqrt{2} y y_i - d\big) v_i + (dy_i - ex_i). \end{aligned} \tag{8}$$

We treat (7) and (8) as equations defining u_i and v_i implicitly as functions of the parameters x, y, d, e, f, with z given by the equation $z = \sqrt{1 + x^2 + y^2}$. Taking derivatives with respect to each parameter $p \in \{x, y, d, e, f\}$ and using the relations $\partial z/\partial x = x/z$ and $\partial z/\partial y = y/z$, we obtain five pairs of equations of the form

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} \dfrac{\partial u_i}{\partial p} \\ \dfrac{\partial v_i}{\partial p} \end{bmatrix} = \begin{bmatrix} P_1^\star \\ P_2^\star \end{bmatrix}$$

from which the partial derivatives $\partial u_i/\partial p$ and $\partial v_i/\partial p$ can be obtained. (The exact expressions for the various parameters are deferred to an appendix.) These partial derivatives are then used analogously as before to iteratively improve estimates for the coefficients of the regression ellipse, this time using $\sum_{i=1}^n ((u_i - x_i)^2 + (v_i - y_i)^2)$ as the criterion for the goodness of fit. To test this algorithm, we choose an example reported in [9], namely that of finding the regression ellipse through the following data points.

x	1	2	5	7	9	3	6	8
y	7	6	8	7	5	7	2	4

In [9], it is reported that 71 iteration steps were required to find the "best ellipse" with respect to the chosen criterion of best fit in the geometric sense. The algorithm presented here performs very favorably in this example; the best fit is reached after about 10 iterations, as is shown in Figure 5.

The above examples suggest that the presented approach works quite well, and it would certainly be worth the effort to systematically compare its performance (as to rate of convergence, robustness or other criteria) with that of other methods. Moreover, if regres-

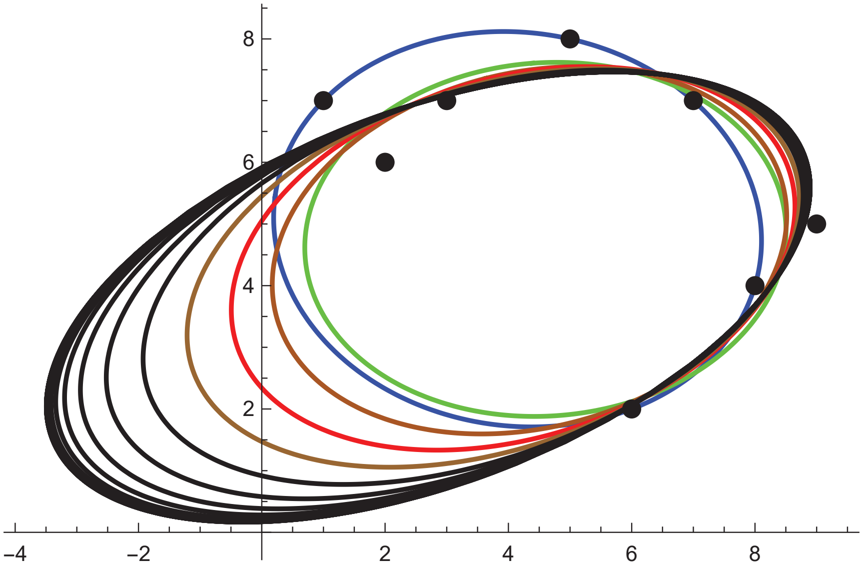

Figure 5. Iterative determination of the ellipse of geometric best fit in the chosen example, with an ellipse through five of the given data points (blue) taken as initial estimate.

sion ellipses are sought for real-world data for which statistical information is available, it would also be of interest to assess the statistical qualities of the parameter estimates obtained with our approach; this leads up to interesting questions concerning the interplay between differential geometry and statistics (cf. [2–4, 7, 12, 13, 18]). However, these issues are beyond the scope of the present paper, which has no higher ambition than to convey how, in a concrete example, geometric ideas can be applied to solve a parameter estimation problem.

The approach presented here requires no deep knowledge of differential geometry, but the grasp of some key concepts (manifold, Riemannian metric, geodesic, direct product of Riemannian manifolds) and the geometric intuition to apply these concepts in a concrete setting. This seems to be the typical way differential geometric ideas enter into engineering disciplines such as robotics, satellite control or image processing. The ever-growing importance of the concept of a differentiable manifold stands in contrast to its subtlety; it is still a didactical challenge to introduce this concept to novices not well-versed in modern topology and differential calculus. The fact that the concept of a manifold is on the one hand quite natural and geometrically appealing, on the other hand quite subtle and not so easy to define precisely, was felt early on in the development of differential geometry. Even in the second edition of his textbook [6], Elie Cartan, one of the masters of the field and an important contributor to the theory, desisted from giving a clear-cut general definition and preferred discussing examples: “La notion générale de variété est assez difficile à definir avec précision.” We feel that examples such as the one presented here do not just serve the purpose of applying differential geometric methods, but can also help to motivate and introduce the relevant concepts and hence are of some didactical value.

Appendix

In this appendix, we list the formulas required to implement the proposed algorithm to determine the ellipse of geometric best fit associated with a given set of data points. The partial derivatives with respect to x are determined by

$$\begin{aligned}
X_{11} &= u_i(x+z) + v_i y + (d/\sqrt{2}),\\
X_{12} &= u_i y + v_i(z-x) + (e/\sqrt{2}),\\
X_{21} &= 2u_i y - 2v_i x - x_i y + y_i(x+z) + (e/\sqrt{2}),\\
X_{22} &= -2u_i x - 2v_i y - x_i(z-x) + y_i y - (d/\sqrt{2}),\\
X_1^\star &= \big(v_i^2(z-x) - u_i^2(z+x)\big)/(2z),\\
X_2^\star &= 2u_i v_i + \big(x_i v_i(x-z) - u_i y_i(x+z)\big)/z.
\end{aligned}$$

The partial derivatives with respect to y are determined by

$$\begin{aligned}
Y_{11} &= u_i(x+z) + v_i y + (d/\sqrt{2}),\\
Y_{12} &= u_i y + v_i(z-x) + (e/\sqrt{2}),\\
Y_{21} &= 2u_i y - 2v_i x - x_i y + y_i(x+z) + (e/\sqrt{2}),\\
Y_{22} &= -2u_i x - 2v_i y - x_i(z-x) + y_i y - (d/\sqrt{2}),\\
Y_1^\star &= -u_i v_i - (u_i^2 + v_i^2)y/(2z),\\
Y_2^\star &= -u_i^2 + v_i^2 + x_i u_i - y_i v_i + (x_i v_i - y_i u_i)y/z.
\end{aligned}$$

The partial derivatives with respect to d are determined by

$$\begin{aligned}
D_{11} &= u_i(x+z) + v_i y + (d/\sqrt{2}),\\
D_{12} &= u_i y + v_i(z-x) + (e/\sqrt{2}),\\
D_{21} &= 2u_i y - 2v_i x - x_i y + y_i(x+z) + (e/\sqrt{2}),\\
D_{22} &= -2u_i x - 2v_i y - x_i(z-x) + y_i y - (d/\sqrt{2}),\\
D_1^\star &= -u_i/\sqrt{2},\\
D_2^\star &= (v_i - y_i)/\sqrt{2}.
\end{aligned}$$

The partial derivatives with respect to e are determined by

$$\begin{aligned}
E_{11} &= u_i(x+z) + v_i y + (d/\sqrt{2}),\\
E_{12} &= u_i y + v_i(z-x) + (e/\sqrt{2}),\\
E_{21} &= 2u_i y - 2v_i x - x_i y + y_i(x+z) + (e/\sqrt{2}),\\
E_{22} &= -2u_i x - 2v_i y - x_i(z-x) + y_i y - (d/\sqrt{2}),\\
E_1^\star &= -v_i/\sqrt{2},\\
E_2^\star &= (x_i - u_i)/\sqrt{2}.
\end{aligned}$$

The partial derivatives with respect to f are determined by

$$
\begin{aligned}
F_{11} &= u_i(x+z) + v_i y + (d/\sqrt{2}),\\
F_{12} &= u_i y + v_i(z-x) + (e/\sqrt{2}),\\
F_{21} &= 2u_i y - 2v_i x - x_i y + y_i(x+z) + (e/\sqrt{2}),\\
F_{22} &= -2u_i x - 2v_i y - x_i(z-x) + y_i y - (d/\sqrt{2}),\\
F_1^\star &= -1/\sqrt{2},\\
F_2^\star &= 0.
\end{aligned}
$$

Funding. Financial support by the Klaus Tschira Foundation for this work is gratefully acknowledged.

References

[1] S. J. Ahn, W. Rauh, and B. Oberdorfer, Least squares fitting of circle and ellipse. In *Mustererkennung 1998; 20. DAGM-Symposium*, pp. 333–340, Springer, Berlin, 1998

[2] S. Amari, *Information geometry and its applications*. Appl. Math. Sci. 194, Springer, Tokyo, 2016

[3] S. Amari and H. Nagaoka, *Methods of information geometry*. Transl. Math. Monogr. 191, American Mathematical Society, Providence; Oxford University Press, Oxford, 2000

[4] N. Ay, J. Jost, H. V. Lê, and L. Schwachhöfer, *Information geometry*. Ergeb. Math. Grenzgeb. (3) 64, Springer, Cham, 2017

[5] D. M. Bates and D. G. Watts, *Nonlinear regression analysis and its applications*. Wiley Ser. Probab. Math. Stat. Appl. Probab. Stat., John Wiley & Sons, New York, 1988

[6] É. Cartan, *Leçons sur la géométrie des espaces de Riemann*. Gauthier-Villars, Paris, 1946

[7] Y. Chikuse, *Statistics on special manifolds*. Lecture Notes in Stat. 174, Springer, New York, 2003

[8] A. W. Fitzgibbon, M. Pilu, and R. B. Fisher, Direct least squares fitting of ellipses. *IEEE Trans. Pattern Anal. Machine Intell.* **21**, no. 5, 476–480 (1999)

[9] W. Gander, G. H. Golub, and R. Strebel, Least-squares fitting of circles and ellipses. *BIT* **34** (1994), no. 4, 558–578

[10] W. Gander, G. H. Golub, and R. Strebel, Least-squares fitting of circles and ellipses. *Bull. Belg. Math. Soc. Simon Stevin Suppl.*, 63–84 (1996)

[11] R. Halíř, J, Flusser: *Numerically stable direct least squares fitting of ellipses*, Proc. 6th Int. Conf. in Central Europe on Computer Graphics and Visualization, Pilsen 1998

[12] V. Patrangenaru and L. Ellingson, *Nonparametric statistics on manifolds and their applications to object data analysis*. CRC Press, Boca Raton, 2016

[13] X. Pennec, Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *J. Math. Imaging Vision* **25** (2006), no. 1, 127–154

[14] A. Ramsay and R. D. Richtmyer, *Introduction to hyperbolic geometry*. Universitext, Springer, New York, 1995

[15] W. F. Reynolds, Hyperbolic geometry on a hyperboloid. *Amer. Math. Monthly* **100** (1993), no. 5, 442–455

[16] K. Spindler, *Höhere Mathematik – Ein Begleiter durch das Studium*, Verlag Harri Deutsch, Frankfurt am Main, 2010

[17] K. Spindler, *Höhere Mathematik – Aufgaben und Lösungen. Vol. 1–3*. Verlag Europa-Lehrmittel, Haan, 2021

[18] G. S. Watson, *Statistics on spheres*. Univ. Arkansas Lect. Notes Math. Sci. 6, John Wiley & Sons, New York, 1983

[19] J. Yu, S. R. Kulkarni, and H. V. Poor, Robust fitting of ellipses and spheroids. In *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, pp. 94–98, IEEE Press, Piscataway, 2010

[20] J. Yu, H. Zheng, S. R. Kulkarni, and H. V. Poor, Outlier elimination for robust ellipse and ellipsoid fitting. In *2009 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 33–36, IEEE Press, Piscataway, 2010

Lilija Naiwert
Kleinaustraße 20
65201 Wiesbaden, Germany
Lilija.Naiwert@googlemail.com

Karlheinz Spindler (corresponding author)
Hochschule RheinMain
Kurt-Schumacher-Ring 18
65197 Wiesbaden, Germany
Karlheinz.Spindler@hs-rm.de