J. Korea Soc. Math. Educ. Ser. B: Pure Appl. Math.
Volume 11, Number 2 (May 2004), Pages 103–115

ISSN 1226-0657

TOTAL LEAST SQUARES FITTING WITH QUADRICS

HELMUTH SPÄTH

ABSTRACT. A computational algorithm is developed for fitting given data in the plane or in 3-space by implicitly defined quadrics. Implicity implies that the type of the quadric is part of the model and need not be known in advance. Starting with some estimate for the coefficients of the quadric the method will alternatively determine the shortest distances from the given points onto the quadric and adapt the coefficients such as to reduce the sum of those squared distances. Numerical examples are given.

1. PROBLEM STATEMENT

Let be given data points

$$(p_i, q_i) \quad \text{or} \quad (p_i, q_i, r_i) \qquad (i = 1, \ldots, m) \tag{1}$$

in the plane or in space that should be fitted by a quadric

$$g_1(x, y) = a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6 = 0 \tag{2}$$

or

$$f_1(x, y, z) = a_1x^2 + a_2y^2 + a_3z^2 + a_4xy + a_5xz + a_6yz + a_7x + a_8y + a_9z + a_{10} = 0, \tag{3}$$

respectively.

If the type of the quadric is known in advance you should better use a parametric representation of it like

$$x = A + P\cos t, \quad y = B + Q\sin t \qquad (0 \le t < 2\pi) \tag{4}$$

Received by the editors January 9, 2004.
2000 *Mathematics Subject Classification.* 65D10, 65C60.
Key words and phrases. total least squares, quadrics.

in the case of an ellipse Ahn, Rauh & Recknagel [1], Ahn, Rauh & Warnecke [2] Gander, Golub & Strebel [4], Späth [11] ($P = Q$: circle Späth [10]), like

$$x = A + P\cos u \sin v, \quad y = B + Q\cos u \cos v, \quad z = C + R\sin u, \tag{5}$$
$$-\pi \le u < \pi,\ -\frac{\pi}{2} \le v < \frac{\pi}{2}$$

in the case of an ellipsoid Späth [12, 15] ($P = Q = R$: sphere Späth [13]). These representations would additionally have to be rotated, *i. e.*, (4) by an unknown elementary rotation in the $x - y$-plane and (5) by three unknown rotations in the $x - y$, $x - z$, and $y - z$ planes.

In order to have the same number of unknowns (center, half axes, rotation angle(s)) both in (4) and (2) and in (5) and (3) and because we can multiply (2) and (3) by some nonzero factor without changing them it turns out that we need some normalization. Excluding quadrics passing through the origin, we set

$$a_6 = 1 \quad \text{in (2) and} \quad a_{10} = 1 \quad \text{in (3), respectively,} \tag{6}$$

to overcome this. Other normalizations are discussed in some references given in Späth [12] and in Späth [15] .

Now if the type of the conic is not known in advance you will have to choose (2) or (3) together with (6). But it is true that in the case of conic sections in the plane you can also use NURBS of degree two Seufer & Späth [6] containing some unknown parameter w whose size determines the type of the quadric (2); it is not clear whether this method could be extended to quadrics in space.

Fitting with (2) or (3) s.t. (6) means to find some vector $\boldsymbol{a} = (a_1, \ldots, a_5)^T$ or $\boldsymbol{a} = (a_1, \ldots, a_9)^T$, respectively, such that the sum of squared shortest distances from the given points (1) onto the quadric is (globally) minimized. Any formulae will be developed only for (3); the equivalent ones for (2) might be got by dropping third components of vectors, *i. e.*, those containing z and r_i, and changing (15).

The computational method to be developed is as follows. In Section 2 we describe at first how to give an initial ($t = 0$) estimate $\boldsymbol{a}^{(t)}$ for the coefficients in (3) using the given data. Then in Section 3 we will use NEWTON's method to determine foot points $(x_i, y_i, z_i) = (x_i(\boldsymbol{a}^{(t)}), y_i(\boldsymbol{a}^{(t)}), z_i(\boldsymbol{a}^{(t)}))$ ($i = 1, \ldots, m$) on the corresponding quadric drawing perpendiculars onto that one with $\boldsymbol{a} = \boldsymbol{a}^{(t)}$. Some heuristic will always empirically give that foot point with the shortest distance. In Section 4 we discuss how to improve $\boldsymbol{a}^{(t)}$ using a damped version of the GAUSS-NEWTON method to reduce the objective function. Using the improved value $\boldsymbol{a}^{(t+1)}$ we start again to look for the optimal foot points and so on. Finally in Section 5 we will give numerical examples and discuss the experiences with the overall method.

2. Starting values

Some starting vector $\boldsymbol{a}^{(0)}$ as required may be generated as follows. From the given data points (p_i, q_i, r_i), $i = 1, \ldots, m$, we calculate

$$\begin{aligned} p_\ell &= \min_i p_i, \quad p_u = \max_i p_i, \\ q_\ell &= \min_i q_i, \quad q_u = \max_i q_i, \\ r_\ell &= \min_i r_i, \quad r_u = \max_i r_i. \end{aligned} \tag{7}$$

The center of the cube with the eight vertices (p_ℓ, q_ℓ, r_ℓ), $(p_\ell, q_\ell, r_u), \ldots,$ (p_u, q_u, r_u) is given by

$$p_m = \frac{p_\ell + p_u}{2}, \; q_m = \frac{q_\ell + q_u}{2}, \; r_m = \frac{r_\ell + r_u}{2}. \tag{8}$$

This cube contains all given data points (1). This is also true for a sphere with the same center (8) and radius

$$R = \sqrt{(p_m - p_\ell)^2 + (q_m - q_\ell)^2 + (r_m - r_\ell)^2}. \tag{9}$$

That sphere as starting quadric, however, may have the disadvantage that all perpendiculars would start from its interior. Thus we reduce R by some factor γ with $0 < \gamma < 1$ - normally $\gamma = .5$ - to get

$$\begin{aligned} a_1^{(0)} &= a_2^{(0)} = a_3^{(0)} = \alpha, \\ a_4^{(0)} &= a_5^{(0)} = a_6^{(0)} = 0, \\ a_7^{(0)} &= -2\alpha p_m, \; a_8^{(0)} = -2\alpha q_m, \; a_9^{(0)} = -2\alpha r_m, \\ \alpha &= 1/(p_m^2 + q_m^2 + r_m^2 - \gamma^2 R^2), \; a_{10}^{(0)} = 1. \end{aligned} \tag{10}$$

These starting values will work in a lot of cases but surely not always.

3. Determining optimal foot points

Assuming $\boldsymbol{a} = \boldsymbol{a}^{(t)}$ to be available in the t-th iteration (starting with (10) for $t = 0$) we consider the problem of finding for some given point $(p, q, r) = (p_i, q_i, r_i)$,

$i = 1, \ldots, m$, the shortest perpendicular onto the current quadric, *i. e.*, we look for $(x, y, z) = (x_i, y_i, z_i)$ on the quadric such that

$$S(x,y,z) = \frac{1}{2}\left[(x-p)^2 + (y-q)^2 + (z-r)^2\right] \tag{11}$$

will globally be minimized s.t. to (3) being valid for $\boldsymbol{a} = \boldsymbol{a}^{(t)}$ and (x, y, z).

S is bounded below and convex. The LAGRANGIAN function is

$$L(x,y,z,\lambda) = S(x,y,z) - \lambda f_1(x,y,z). \tag{12}$$

The necessary conditions for some minimum, *i. e.*,

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} = \frac{\partial L}{\partial z} = 0 \tag{13}$$

give in turn

$$\begin{aligned} (x-p) - \lambda\,(2a_1x + a_4y + a_5z + a_7) &= 0, \\ (y-q) - \lambda\,(2a_2y + a_4x + a_6z + a_8) &= 0, \\ (z-r) - \lambda\,(2a_3z + a_5x + a_6y + a_9) &= 0. \end{aligned} \tag{14}$$

Eliminating λ *e. g.*, from the first equation and putting it into the two other ones gives

$$\begin{aligned} f_2(x,y,z) &= (y-q)(2a_1x+a_4y+a_5z+a_7) - (x-p)(2a_2y+a_4x+a_6z+a_8) = 0, \\ f_3(x,y,z) &= (z-r)(2a_1x+a_4y+a_5z+a_7) - (x-p)(2a_3z+a_5x+a_6y+a_9) = 0. \end{aligned} \tag{15}$$

Together with $f_1(x, y, z) = 0$ from (3) we have three nonlinear equations for the unknowns (x, y, z). For geometrical reasons it is clear that these equations will have at least one solution. Indeed the number of solutions is known for special quadrics. We normally have two solutions for a sphere Späth [14], one or three for a rotated paraboloid Späth [11], and two up to six for rotated ellipsoids Späth [13] and hyperboloids. (In the case of plane quadrics, *i. e.*, conic sections, those numbers would be two, one or three, and two up to four Späth [10, 11, 12]). In the (normal) case of several solutions we have to determine all of them and pick up that one giving the shortest distance, *i. e.*, the global minimum.

To afford this we used a damped version of NEWTON's method implemented as FORTRAN subroutine TAYLOR in Späth [7, 8]. The necessary partial derivatives

$$\frac{\partial f_k}{\partial x}, \; \frac{\partial f_k}{\partial y}, \; \frac{\partial f_k}{\partial z} \qquad (k = 1, 2, 3) \tag{16}$$

could easily be calculated and are linear functions of x, y, z. Because of linearity forward or central divided differences – an option in TAYLOR – give exactly the same results. To make it easier we used this fact in the following. As starting values we defined (for each $i = 1, \dots, m$) N different ones, namely

$$\begin{aligned} x_i^{(0)} &= p_\ell + v(p_u - p_\ell), \\ y_i^{(0)} &= q_\ell + v(q_u - q_\ell), \\ z_i^{(0)} &= r_\ell + v(r_u - r_\ell), \end{aligned} \tag{17}$$

where v means at each appearance some new (equally in $[0,1]$ distributed) pseudo-random number. Empirically in each of hundreds of thousand iterations NEWTON's method with those starting values converged within 4 and 7 iterations without any damping to some solution correct to five decimal digits.

We have a maximum of six possible solutions. Thus if we get for N starting values (for each $i = 1, \dots, m$) six solutions we could easily select the global minimum of the N found solutions. For our examples $N = 50$ was sufficiently large. Even if this strategy would not work the overall method does not necessarily break down.

4. Adapting the coefficients of the current quadric

Let be $\big(x_i = x_i(\boldsymbol{a}),\ y_i = y_i(\boldsymbol{a}),\ z_i = z_i(\boldsymbol{a})\big)$, $i = 1, \dots, m$, the optimal foot points on the current quadric (3) with $\boldsymbol{a} = \boldsymbol{a}^{(t)}$. Introducing the vector

$$\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{a}) = (x_1 - p_1, y_1 - q_1, z_1 - r_1, \dots, x_m - p_m, y_m - q_m, z_m - r_m)^T \tag{18}$$

of length $3m$ our overall objective can be written as

$$\frac{1}{2}\,\|\boldsymbol{s}(\boldsymbol{a})\|^2 = \frac{1}{2}\,\boldsymbol{s}(\boldsymbol{a})^T \boldsymbol{s}(\boldsymbol{a}). \tag{19}$$

We now try to improve the current $\boldsymbol{a} = \boldsymbol{a}^{(t)}$ by $\Delta\boldsymbol{a} = \Delta\boldsymbol{a}^{(t)}$ using the one term TAYLOR's expansion

$$\boldsymbol{s}(\boldsymbol{a} + \Delta\boldsymbol{a}) \approx \boldsymbol{s}(\boldsymbol{a}) + \boldsymbol{J}(\boldsymbol{a})\Delta\boldsymbol{a}, \tag{20}$$

where

$$\boldsymbol{J}(\boldsymbol{a}) = \left(\frac{\partial s_i}{\partial a_k}\right)_{i=1,\dots,3m;\, k=1,\dots,9} \tag{21}$$

is the Jacobian, by minimizing the Euclidean norm of the right-hand side of (20) w.r.t. $\Delta\boldsymbol{a}$. Equivalent to this is to solve the overdetermined system of linear equations

$$\boldsymbol{J}(\boldsymbol{a})\Delta\boldsymbol{a} = -\boldsymbol{s}(\boldsymbol{a}) \tag{22}$$

in the least squares sense. This can be done *e. g.*, by the modified GRAM-SCHMIDT method. We used the corresponding FORTRAN subroutine MGS from Späth [9].

The rest of this well-known GAUSS-NEWTON method is to find a value for some step size control parameter β such that (19) decreases for

$$\boldsymbol{a}^{(t+1)} = \boldsymbol{a}^{(t)} + \beta \Delta \boldsymbol{a}^{(t)}. \tag{23}$$

If GAUSS-NEWTON is successful in this sense, then we go back to Section 3 and so on. In this case this alternating method is a descent method. But determining such a β may not be able in some cases for special starting values.

Calculating the least squares solution of (22) requests estimating $\boldsymbol{J}(\boldsymbol{a})$ that is more explicitly given by its k-th column ($k = 1, \ldots, 9$) of length $3m$, namely by

$$\left(\frac{\partial x_1}{\partial a_k}, \frac{\partial y_1}{\partial a_k}, \frac{\partial z_1}{\partial a_k}, \ldots, \frac{\partial x_m}{\partial a_k}, \frac{\partial y_m}{\partial a_k}, \frac{\partial z_m}{\partial a_k} \right)^T. \tag{24}$$

Closed but complicated expressions for (24) are given in Ahn, Rauh & Warnecke [2], Gulliksson, Söderkvist & Watson [5]. We prefer to use central divided differences for each of the m triples in (24), *i. e.*,

$$\begin{aligned} \frac{\partial x_i}{\partial a_k} &\approx \frac{x_i(\boldsymbol{a}^+) - x_i(\boldsymbol{a}^-)}{2h_k}, \\ \frac{\partial y_i}{\partial a_k} &\approx \frac{y_i(\boldsymbol{a}^+) - y_i(\boldsymbol{a}^-)}{2h_k}, \quad (i = 1, \ldots, m;\ k = 1, \ldots, 9) \\ \frac{\partial z_i}{\partial a_k} &\approx \frac{z_i(\boldsymbol{a}^+) - z_i(\boldsymbol{a}^-)}{2h_k}, \end{aligned} \tag{25}$$

where

$$\begin{aligned} \boldsymbol{a}^+ &= (a_1, \ldots, a_k + h_k, \ldots, a_9), \\ \boldsymbol{a}^- &= (a_1, \ldots, a_k - h_k, \ldots, a_9), \end{aligned} \tag{26}$$

and where h_k are suitable step sizes, *e. g.*, Späth [8].

$$h_k = \begin{cases} .0001 & \text{if } |a_k| \le .01, \\ .01|a_k| & \text{if } |a_k| > .01. \end{cases} \tag{27}$$

It is true that in this way we will have for each of the m triples (x_i, y_i, z_i) two foot point problems, *i. e.*, $2m$ additional ones, namely for $\boldsymbol{a}^+$ and $\boldsymbol{a}^-$ changing with k, but those are very easy to solve numerically.

Using the actual foot points $\big(x_i(\boldsymbol{a}), y_i(\boldsymbol{a}), z_i(\boldsymbol{a})\big)$ as starting values it normally takes not more than two NEWTON iterations to find $\big(x_i(\boldsymbol{a}^+), y_i(\boldsymbol{a}^+), z_i(\boldsymbol{a}^+)\big)$ and $\big(x_i(\boldsymbol{a}^-), y_i(\boldsymbol{a}^-), z_i(\boldsymbol{a}^-)\big)$ needed in (23). For the convergence of the overall method it is essential to repeat this until getting some suitable β (if possible at all) in (23) to decrease the objective.

5. Numerical examples and conclusions

In order to demonstrate the capability of the described algorithm we will give five examples. The first four examples are given for quadrics in the plane. They show that the type of the conic section will be identified. The fifth example uses data in 3-space.

As starting values for $\boldsymbol{a}^{(0)}$ we used (10) and $N = 50$ for (17). Instead of the original given data (1) we sometimes also used

$$(p_i - \overline{p}, q_i - \overline{q}) \quad \text{or} \quad (p_i - \overline{p}, q_i - \overline{q}, r_i - \overline{r}) \qquad (i = 1, \ldots, m), \tag{28}$$

where $\overline{p}, \overline{q}$, and $\overline{r}$ are the means of the corresponding components. The reason is that occasionally the convergence behaviour will be improved in this way. The value for S must be the same as for the original data, but, of course, the coefficients $\boldsymbol{a}$ are different.

Example 1. The data points Späth [15] were

p	1	3	4	5	6	4	2	0	-1	-2	-1
q	-2	-2	0	1	4	5	4	4	2	-1	-3

The starting values (10) were

$$\boldsymbol{a}^{(0)} = (-.33333, -.33333, .00000, 1.33333, .66667)^T.$$

Within 5 iterations the method converged to

$$\boldsymbol{a}^{(5)} = (-.10968, -.11711, .12641, .25360, .04218)^T$$

with $S = .71561$. The same solution was also found for

$$a^{(0)} = (.1, .2, .3, .4, .5)^T$$

and for

$$a^{(0)} = (-.0571, -.1268, -.0714, -, 0143, -.0571)^T.$$

The solution (an ellipse) together with the given data points can be seen in Figure 1. Data for drawing the figures were generated by varying y within $[q_\ell, q_u]$ and solving the corresponding quadric equations (2) for x using the found $\boldsymbol{a}$.

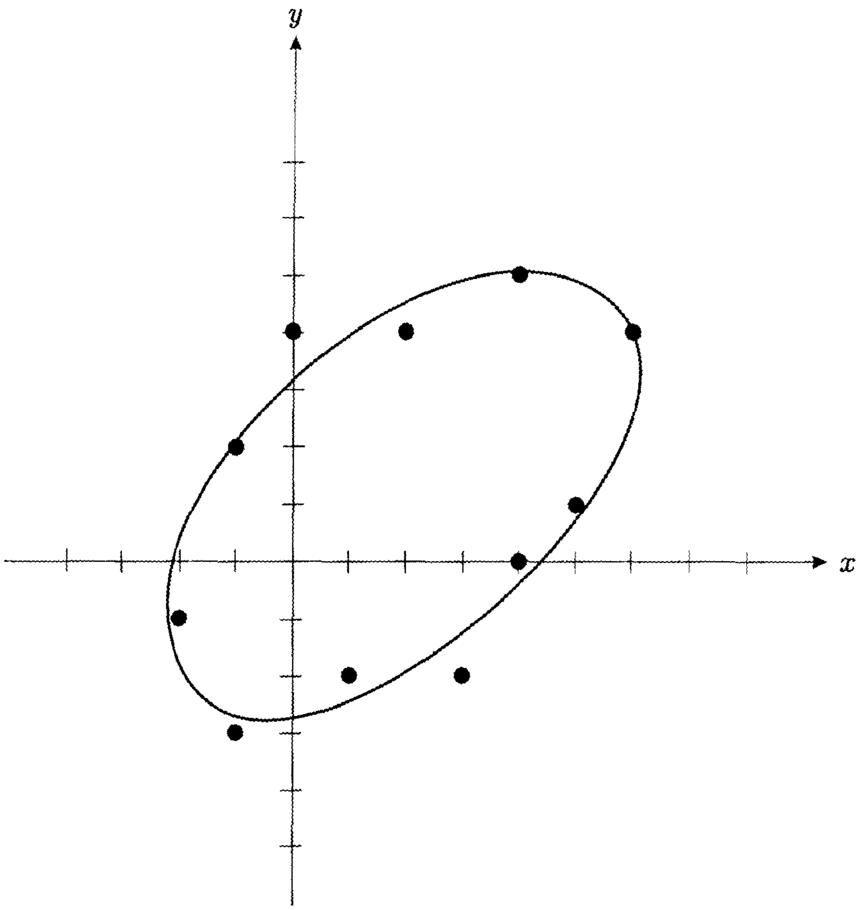

Figure 1. Ellipse 1

Example 2. The data used in Späth [15] were

p	1	2	3	5	7	9	8	6
q	7	6	7	8	7	5	4	2

They are slightly modified from an example in Gander, Golub & Strebel [4]. It is not clear in advance that they should indicate an ellipse. For the starting value (10) in the case of the transformed data (28), *i. e.*,

$$\boldsymbol{a}^{(0)} = (-.17631, -.17631, .00000, -.04408, -.26446)^T$$

we got after 13 iterations

$$\boldsymbol{a}^{(13)} = (-.04910, -.14120, .08054, -.08147, -.35010)^T$$

with $S = 0.6867$. The resulting ellipse is in Figure 2.

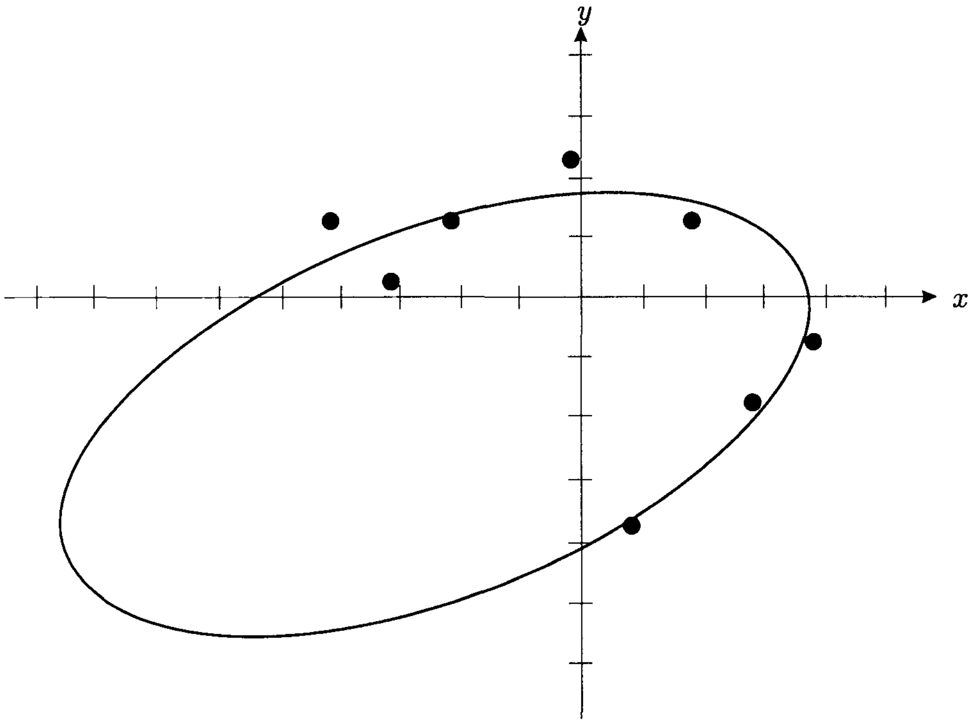

Figure 2. Ellipse 2

Example 3. The data

p	1	2	3	5	7	5	9	4
q	4	2	6	6	-1	0	8	2

were used in Späth [12] to demonstrate the possibility of fitting with one branch of a hyperbola. Using again the starting value (10), *i. e.*,

$$\boldsymbol{a}^{(0)} = (.03548, .03548, .00000, -.35477, -.27834)^T$$

we got after 7 iterations

$$\boldsymbol{a}^{(7)} = (-.01250, .05309, .01776, -.11131, -.45047)^T$$

with $S = .58063$. Applying the transformed data (28) and

$$\boldsymbol{a}^{(0)} = (-.11368, -.11368, .00000, .11368, .02842)^T$$

via (10) we got

$$\boldsymbol{a}^{(5)} = (.03126, -.13276, -.04438, .40975, .03047)^T$$

and the same value for S. See Figure 3 for the resulting hyperbola.

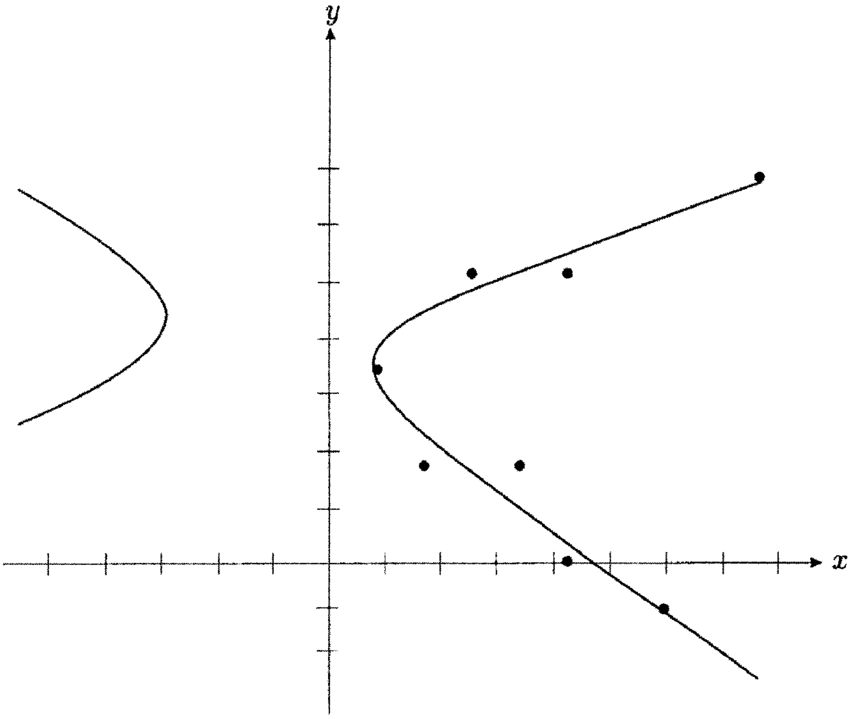

Figure 3. Hyperbola 1

Example 4. The data points

p	-7	-3	0	1	1	0
q	9	5	4	3	5	8

from Späth [11] either indicate a hyperbola or a parabola. Again for the starting values (10), *i. e.*,

$$\boldsymbol{a}^{(0)} = (.02581, .02581, .00000, .15484, -.30968)^T$$

we got

$$\boldsymbol{a}^{(12)} = (.06265, .03678, .10221, -.43768, -.41637)^T$$

with $S = .23182$. See Figure 4 for the resulting parabola.

Example 5. At first we generated $m = 10$ data points

p	1	-3	-1	-3	-5	1	3	1	-1	1
q	1	-3	-1	3	-1	-1	-3	-1	5	-5
r	5	3	5	-3	1	5	3	-5	1	-1

phere with the origin as center and radius $\sqrt{27}$.

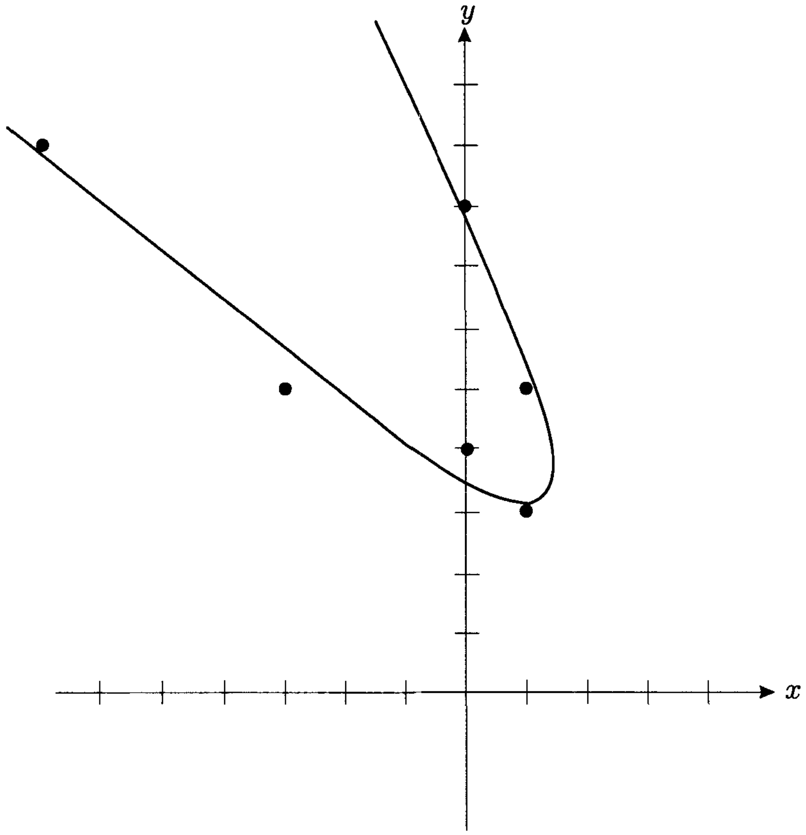

Figure 4. Parabola 1

Then we heavily disturbed each number by up to $\pm.8$ and made a translation by the vector $(1, 2, 3)^T$. This resulted as

p	2.8	-2.5	-.6	-2.5	-4.5	2.3	4.4	2.3	-.4	1.8
q	3.6	-.5	.8	4.1	.5	.5	-.6	-1.3	7.3	-3.7
r	7.6	6.5	8.4	.3	.4	7.5	6.5	-2.4	3.5	1.6

Starting with (10), *i. e.*,

$$\boldsymbol{a}^{(0)} = (-.13226, -.13226, -.13226, 0, 0, 0, -.01323, .47615, .79358)^T$$

we got after 7 iterations

$$\boldsymbol{a}^{(7)}$$
$$= (-.05898, -.10846, -.09854, -.09109, -.07166, .05492, .54158, .30874, .63406)^T.$$

As there are just 10 data points and 9 parameters to be found $S \approx 0$ must be expected. Indeed we got $S = .014361$. Deleting one point, *e. g.*, the last one, you should expect $S = 0$. Indeed we arrived at $S = .12365 \times 10^{-9}$.

As expected in all cases we received the evidently global minimum using the starting values (10). Though a lot of foot point problems had to be solved in the above examples, the overall computing time was just some seconds on a PC.

REFERENCES

1. S. J. Ahn, W. Rauh & M. Recknagel: Geometric fitting of line, plane, circle, sphere, and ellipse. *ABW-Workshop 6, Technische Akademie Esslingen* (pp. 25–26), 1999.
2. S. J. Ahn, W. Rauh, and H. J. Warnecke: Best-fit of implicit surfaces and plane curves. In: T. Lyche & L. L. Schumaker (Eds.), *Mathematical methods for curves and surfaces (Oslo, 2000)* (pp. 1–14). Innov. Appl. Math., Vanderbilt Univ. Press, Nashville, TN, 2001. CMP1857819
3. ______: Least-squares orthogonal distances fitting of circle, sphere, ellipse, hyperbola, and parabola. *Pattern Recognition*, **34** (2001), 2283–2303.
4. W. Gander, G. H. Golub & R. Strebel: Least-squares fitting of circles and ellipses. *BIT* **34** (1994), no. 4, 558–578. MR **97j**:65020
5. M. Gulliksson, I. Söderkvist & G. A. Watson: Implicit surface fitting using directional constraints. *BIT* **41** (2001), no. 2, 308–321. MR **2002c**:65023
6. I. Seufer & H. Späth: Least squares fitting of conic sections with type identification by NURBS of degree two. *Math. Commun.* **4** (1999), no. 2, 207–218. CMP1746383
7. H. Späth: The damped taylor's series method for minimizing a sum of squares and for solving systems of nonlinear equations. *Collected Algorithms of the ACM* (pp. 315–R1), 1967.
8. ______: *Algorithmen fur multivariable Ausgleichsmodelle.* Verfahren der Datenverarbeitung. R. Oldenbourg Verlag, Munich-Vienna, 1974. MR **53**#14864
9. ______: *Mathematical algorithms for linear regression. Translated and revised from the 1987 German original by the author.* Computer Science and Scientific Computing. Academic Press, Inc., Boston, MA, 1992. MR **93a**:62002
10. ______: Least-squares fitting by circles. *Computing* **57** (1996), no. 2, 179–185. MR **97e**:65019
11. ______: Orthogonal squared distance fitting with parabolas. In: G. Alefeld & J. Herzberger (Eds.), *Proceedings of the IMACS-GAMM International Symposium on Numerical methods and error bounds (Oldenburg, 1995)* (pp. 261–269). Akademie Verlag, Berlin, 1996. MR **98c**:65016

12. ______: Least-squares fitting of ellipses and hyperbolas. *Comput. Statist.* **12** (1997), no. 3, 329–341. CMP1477268
13. ______: Orthogonal least squares fitting by conic sections. *Recent advances in total least squares techniques and errors-in-variables modeling (Leuven, 1996)* (pp. 259–264). SIAM, Philadelphia, PA, 1997. CMP1447482
14. ______: Least-square fitting with spheres. *J. Optim. Theory Appl.* **96** (1998), no. 1, 191–199. MR **98m:**65065
15. ______: Estimating the parameters of an ellipse when angular differences are known. *Computational Statistics* **14** (1999), 491–500.
16. D. A. Turner, I. J. Anderson, J. C. Mason & M. G. Cox: *An algorithm for fitting an ellipsoid to data.* 1999. Unpublished manuscript (to be found in www).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OLDENBURG, 26111 OLDENBURG, GERMANY
Email address: `spaeth@mathematik.uni-oldenburg.de`