2012

Geometric Fitting in Error-In-Variables Model

Qizhuo Huang
University of Alabama at Birmingham

GEOMETRIC FITTING IN ERROR-IN-VARIABLES MODEL

by

QIZHUO HUANG

NIKOLAI CHERNOV, COMMITTEE CHAIR
WEI-SHEN HSIA
CHARLES KATHOLI
IAN KNOWLES
BORIS KUNIN

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama,
The University of Alabama at Birmingham, and The University of Alabama in
Huntsville in partial fulfillment of the requirements for the degree
of Doctor of Philosophy

BIRMINGHAM, ALABAMA

2012

ABSTRACT

GEOMETRIC FITTING IN ERROR-IN-VARIABLE MODEL

QIZHUO HUANG

APPLIED MATHEMATICS

This dissertation is devoted to the study of a popular regression model: Error-In-Variable Model, which has been commonly recognized as one of the key components of computer vision research. In EIV model, a set of data points whose x, y coordinates are subject to random errors is fitted by some geometric shapes such as lines, circles and ellipses. The geometric fitting which minimizes the sum of orthogonal distances from points to geometric shapes is universally recognized as the most desirable solution of the fitting problem. However, there is no explicit form of the solution for nonlinear models (circles, ellipses etc). The problem of fitting circles has been investigated intensively over past a few decades and all major issue appeared to be resolved. Our analysis will focus on a more sophisticated model - fitting ellipses to a set of points.

We will address the issues of existence and uniqueness of the best fitting solution, study the parameter space of all quadratic curves and properties of the objective function and show some peculiar feature of the estimates of geometric parameters for the best fitting ellipse: they have no finite moments.

Our results promote understanding of why computer algorithms keep diverging, return nonsense or crash altogether and help development of more robust, efficient fitting schemes.

Keywords: errors in variables (EIV) models, image processing, geometric fitting, ellipse fitting.

DEDICATION

TO MY BELOVED PARENTS

ACKNOWLEDGEMENTS

Over the past four years I have received support and encouragement from my advisor Professor Nikolai Chernov. His guidance has made this a thoughtful and rewarding journey. I would like to show my gratitude to all my dissertation committee of W. Hsia, C. Katholi, I. Knowles, B. Kunin for their support over the past two years as I moved from an idea to a completed study. Thanks to the University of Alabama at Birmingham awarding me a scholarship, providing me with the financial means to complete this project.

To the faculty of the U.A.B.'s mathematics department, in particular, M. Nkashama, R.Weikard, Y. Karpeshina, I thank you for your continuous support and encouragement. Finally, thanks to my parents, and numerous friends who endured this long process with me, always offering support and love.

Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

This dissertation investigates certain theoretic aspects of a popular regression problem: fitting geometric contours (ellipses, hyperbolas, parabolas, etc,) to a set of observed points which are measured imprecisely in both coordinates. This topic is known as Error-in-Variables (EIV) model. It is fundamentally different and much more complicated than the classical regression model which assumes that only one variable is subject to random error (usually called response variable) while the other variable (independent variable) is fixed.

The EIV regression model has been investigated by statisticians since the 1930s [**17, 18**] and its importance has been recognized in many fields such as econometrics, engineering science and image processing. The simplest model of fitting straight lines to a set of observed points dates back to the 1870s [**3, 4, 22**]. All major problems in linear EIV model were resolved by the late 1990s and much attention has been given to nonlinear regression models (circle, ellipse etc) [**5, 7, 29**]. Fitting nonlinear models to data with errors can be divided into two parts. In the first one, the main goal is to approximate data points by a nonlinear function such as a polynomial or an exponential function. The x and y are measured based on different units and thus their errors may have different magnitude (see [**10, 11**] for detail). Second type of nonlinear regression problem assumes both x and y variables are measured in the same units and the choice of the coordinates system is completely arbitrary. Thus the magnitude of errors in both variables are the same. This type of problem commonly arises in the image processing where one often fits the geometric shape to data points on $2D$ image. This dissertation will focus on geometric fitting used in the latter one. The main frame of this regression problem can be formulted as follows:

Suppose one observes n experitmental points $(x_1, y_1), \ldots, (x_n, y_n)$, which are assumed random perturbations of some true points $(\widetilde{x}_i, \widetilde{y}_i)$.

$$x_i = \widetilde{x}_i + \delta_i \quad y_i = \widetilde{y}_i + \varepsilon_i \quad i = 1, \ldots, n \tag{1.1}$$

It is also assumed that all true points $(\widetilde{x}_i, \widetilde{y}_i)$ $i = 1, \ldots, n$ all belongs to an unknown geometric shape, i.e. $P(\widetilde{x}_i, \widetilde{y}_i|\Theta) = 0$ (i=1,. . . ,n) where Θ is the unknown parameter vector. The goal is to find an estimate of Θ so that the geometric shape represented by $y = P(x, y, |\Theta)$ approximates the observed points the best. As a standard assumption in the EIV literature, random errors δ_i's and ε_i's are considered as independently distributed normal random variables with zero mean:

$$\delta_i \sim N(0, \sigma_x^2) \quad \varepsilon_i \sim N(0, \sigma_y^2) \tag{1.2}$$

We can also make the following assumptions about the true points $(\widetilde{x}_i, \widetilde{y}_i)$'s.

First, we can treat the true points $(\widetilde{x}_i, \widetilde{y}_i)$ $(i = 1, \ldots, n)$ as fixed parameters whose values are normally of little interest in the fitting problem. This type of assumption is known as the *functional model*. Or they can be regarded as realization of some underlying random variables such as $N(\mu, \sigma^2)$. Then μ and σ^2 are considered as parameters to be estimated along with the parameters of interest. Such a treatment is known as *structure model* [**20, 21**]. The functional model has been intensively studied and used in real application, especially in image processing. Therefore, this model is adopted throughout this dissertation. We will turn to introduce the most reliable fitting method for the EIV model.

1.1. Geometric fit

In EIV model, there are two approaches we can use to find a circle or ellipse which best fits our data: algebraic fitting and geometric fitting. The geometric fit which minimizes the sum of squares of orthogonal distances from points to the curve is commonly regarded as being more accurate than algebraic fits. It has many nice features:

- It is invariant under translations, rotations, and scaling, i.e., the fitted geometric fitting does not depend on the choice of the coordinate system.
- It coincides with the maximum likelihood estimate of the parameters of the fitted geometric shape under standard statistical assumptions.
- Geometric fit sets a standard for testing the data processing software for coordinate metrology [**1**]

Given some points $P_1, \ldots, P_n \in \mathbb{R}^2$, the objective function is the sum of squares of the distances to a model object (in our case, conic) S:

$$\mathcal{F}(S) = \sum_{i=1}^{n} \big[\operatorname{dist}(P_i, S)\big]^2, \tag{1.3}$$

The objective function depends on the points $P_1, \ldots, P_n$, but in practical settings those are fixed, so the only variable is S, which is regarded as the sole argument of $\mathcal{F}$.

In the case of fitting line $Ax + By + C = 0$ to the data points, one can easily show that

$$\mathcal{F}(A, B, C) = \frac{1}{A^2 + B^2} \sum_{1}^{n} (Ax_i + By_i + C)^2 \tag{1.4}$$

By setting $A^2 + B^2 = 1$, one can find the minimzer for (1.6) explicitly(see section 3.1).

Suppose one try to fit a circle to the points. Then the distance between the circle and each point are

$$d_i = \sqrt{(x_i - a)^2 + (y_i - b)^2} - R \tag{1.5}$$

where (a, b) denotes the center and R the radius of the circle. However, the minimization of the objective function

$$\mathcal{F}(A, B, C) = \sum_{1}^{n} [\sqrt{(x_i - a)^2 + (y_i - b)^2} - R] \tag{1.6}$$

has no closed form of solution. Its solution can only be approximated by iterative algorithms such as Gauss-Newton or Levenberg-Marquard scheme which usually takes

dozens or hundreds of iterations to converge, and there is a chance that they may diverge, return nonsense or crash altogether [**15**].

The situation in ellipse fitting becomes worse. The orthogonal distance between the ellipse and a given point do not even have simple analytic formula. The distance can be computed by equivalently solving a polynomial equation of degree 4, which again requires numerical schemes. The lack of explicit form of solution makes the analytic investigation of the nonlinear model much more difficult. In this dissertation, we will address several fundamental issues that will help promote understanding of the geometric fitting in nonlinear case especially ellipse fitting.

1.2. Organization of the Thesis

Our analysis mainly focus on the following issues:

- Does the best fit always exists? Meaning: does the objective function $\mathcal{F}$ always have a minimum?
- Is the solution always unique? (Meaning: is the minimum of the objective function always unique?)
- The topological properties of the parameter space of quadratic curves and how the objective function behaves in the parameter space
- The moments of estimates of geometric parameters.

The dissertation is organized as follows. The second chapter discuss the issue of existence of the best fitting solution in a general sense. Our approach applies not just to the model collection of ellipses but arbitrary closed sets. The third chapter first reviews the issue of uniqueness of best fitting line and circle. And then provide an example of multiple best fitting ellipse with a computer assisted proof. In chapter 4, we study the algebraic parameter space of quadratic curves confined to the unit sphere in a topological manner. The unit sphere is divided into several domains based on types of curve. We will analyse the topological properties of each domain both seperately and together. Chapter 5 discusses important properties of the objective

function such as continuity and differentiability. In the last chapter, we develop a general strategy for checking infinite moments for geometric parameters and prove that the geometric parameters of the best fitting ellipse do not have finite moments.

1.3. Special Remark

The first few chapters of my dissertation may appear quite similar to those of Ali and especially Hui Ma (who defended her thesis in May). This is because they all worked on the same general topic of fitting circles and ellipses, and in the first few chapters the topic is introduced, with all definitions and general constructions which are about the same for all of whole group.

At the same time their results are all different and there is no single joint result. Hui Ma has mentioned that some theorems and facts are worked out by me and I also quote her results in some of my sections. We all tried to be very discrete in this respect.

The last chapter (Chapter 6) in my dissertation is very special, though. There is nothing like it in Ali's or Hui's theses. It is about infinite moments of the ellipse parameter estimators (center and axes) which might be considered as the most mathematically interesting.

CHAPTER 2

EXISTENCE OF THE BEST FITTING SET

In this chapter we will investigate the problem of existence of the best fitting curve (also see [**30**]). To make a more general discussion, we will deal with all closed sets in $\mathbb{R}^2$, not just some popular models (line, circle, ellipses) used in practical applications. We will develop a general approach to the study of existence of the best fit and we want to know if the objective function representing the sum of squares of orthogonal distances could always achieve its minimum. Also another closely related question about uniqueness will be discussed separately in the next chapter.

The problems of existence and uniqueness of the best fit are often ignored in the real application as the chance that the best fit does not exist is not quite noticeable. If they come up, one either assumes that the best fit exists and is unique, or just points out examples to the contrary without deep investigation. However the investigation might help understand why the computer algorithm fails to find the best solution (diverge or crahes).

We will begin by introducing some basic notations in section 2.1. Then sections 2.2 to 2.6 will provide a theoretical analysis for the issue of best fit with some main theorems. The discussion will involve concepts of continuity and compactness, which we will engage also later in the section 2.8 to treat the models of ellipses and all quadratic curves. Section 2.7 handles the problem of existence by a different approach.

2.1. Distances (review)

Since our fitting problem involves minimization of sum of squares of points to the model object, let us begin by reviewing some necessary distance definition.

Distance between points:

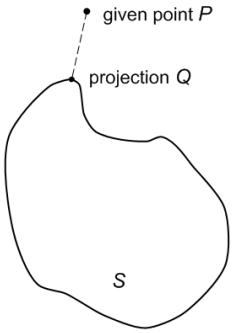


Figure 2.1: Distance from point to closed set

For any two given points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, the standard "geometric" (or "Euclidean") distance is computed by

$$\text{dist}(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \tag{2.1}$$

Distance from point to set

Given a point P and a set $S \subset \mathbb{R}^2$, the geometric distance from P to S is naturally defined by

$$\text{dist}(P, S) = \inf_{Q \in S} \text{dist}(P, Q), \tag{2.2}$$

When the set S is *closed*, the minimum take the place of the infimum (see proof in Appendix), which provides a more practically convenient definition. If such a minimum is attained, there exists a closest point $Q \in S$ to the point P such that

$$\text{dist}(P, S) = \text{dist}(P, Q) = \min_{Q \in S} \text{dist}(P, Q), \tag{2.3}$$

All model objects that are usually fitted to given points - lines, circles, ellipses and other conics - are closed sets.

In most practical cases, the distance from a point P to a set S is obtained by projecting P onto S; then Q is the called the footpoint of the projection. See illustration in the figure 2.1.

As one needs to use orthogonal projection, the distance from P to S is often called *orthogonal distance*.

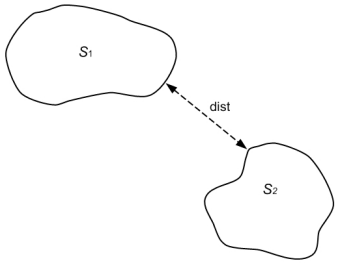


Figure 2.2: Distance from set to set

(Shortest) distance from set to set

Given two sets $S_1, S_2 \subset \mathbb{R}^2$, the distance between S_1 and S_2 is defined by

$$\text{dist}(S_1, S_2) = \inf_{P_1 \in S_1, P_2 \in S_2} \text{dist}(P_1, P_2), \tag{2.4}$$

which is the *shortest* distance from S_1 to S_2. See illustration in Figure 2.2.

The infimum in (2.4) may not be replaced by a minimum even if both sets S_1 and S_2 are closed. For example, let $S_1 = \{(x, y)\colon y = 0\}$ be a straight line (the x axis) and $S_2 = \{(x, y)\colon xy = 1\}$ be a hyperbola whose asymptotes are the x and y axes. The distance between these sets is zero, i.e., $\text{dist}(S_1, S_2) = 0$, but there are no points $P_1 \in S_1$ and $P_2 \in S_2$ such that $\text{dist}(P_1, P_2) = 0$.

However, if one set (say, S_1) is closed and the other (S_2) is compact, the infimum in (2.4) can always be replaced by a minimum (see proof in Appendix). In that case there are closest points $P_1 \in S_1$ and $P_2 \in S_2$ such that $\text{dist}(S_1, S_2) = \text{dist}(P_1, P_2)$. Note that circles and ellipses are closed and bounded, i.e., compact. On the other hand, lines and hyperbolas are closed but not bounded.

Hausdorff distance from set to set

The shortest distance between two sets may be small, but the sets may be overall very different from each other. To describe how far two sets are from each other (or the dissimilarity of two shapes), we can use Hausdorff distance. Given two sets

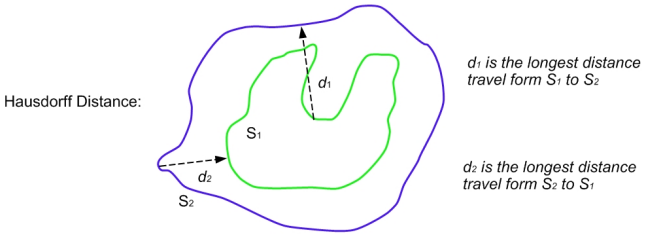


Figure 2.3: Hausdorff distance between two sets

$S_1, S_2 \subset \mathbb{R}^2$, the Hausdorff distance between S_1 and S_2 is defined by

$$\text{dist}_{\text{H}}(S_1, S_2) = \max\Big\{ \sup_{P_1 \in S_1} \text{dist}(P_1, S_2),\ \sup_{P_2 \in S_2} \text{dist}(P_2, S_1) \Big\}, \tag{2.5}$$

which is the longest distance you have to travel if you need to move from one set to the other or vice versa;

If two sets are closed and the Hausdorff distance between them is zero, i.e., $\text{dist}_{\text{H}}(S_1, S_2) = 0$, then they coincide: $S_1 = S_2$. If the Hausdorff distance is small, the two sets nearly coincide with each other. When one set is closed and the other compact (or both are compact), the suprema in (2.5) can be more conveniently replaced by maxima:

$$\text{dist}_{\text{H}}(S_1, S_2) = \max\Big\{ \max_{P_1 \in S_1} \text{dist}(P_1, S_2),\ \max_{P_2 \in S_2} \text{dist}(P_2, S_1) \Big\}, \tag{2.6}$$

This completes our review of standard definitions of distances. Now we are ready to introduce an important concept in the next section .

2.2. Convergence of sequences of sets

In this section we will introduce the notion of convergence for sequences of sets which will involve some type of "distance". Geometrically, a sequence of sets S_i $(i = 1, \ldots)$ converges to limit set S if they become indistinguishable from S. We will make this intuition mathematically rigorous. As we see in the last section, Hausdorff

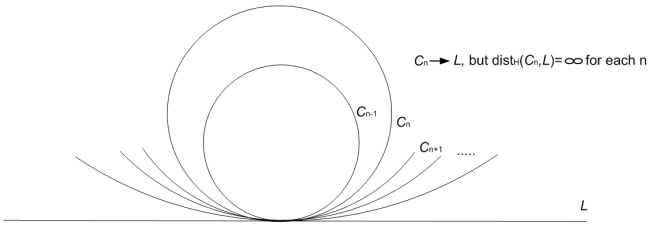


Figure 2.4: Convergence of a sequence of circles

distance set a standard for measuring the "closedness"of two closed sets. However it is only useful for compact sets.

Motivating example

Let us consider a sequence of Circles $C_n = \{x^2 + (y - R)^2 = R^2\}$ passing through the origin $(0, 0)$ and their radius R increases as n grows. Naturally, we would consider this sequence as convergent: it converge to the line L: $y = 0$ as $n \to \infty$.

However, the Hausdorff distance between C_n and L is *infinite*, i.e., $\text{dist}_{\text{H}}(C_n, L) = \infty$ for every n. If you travel away from the point $(0, x)$ on L to C_n, the distance can be arbitrarily large. So the normal Hausdorff distance fails to characterize the "closedness" between two sets even if they are indeed close.

Window-restricted Hausdorff distance Geometrically, we cannot see the whole line L and circle C_n as n gets large but only the parts restricted to a certain finite area move close to that of L, like in the above illustration. Suppose we see objects in some rectangle

$$R = \{-A \le x \le A, \quad -B \le y \le B\}, \tag{2.7}$$

which for the moment will play the role of our "window" through which we look at the plane. Then we see segments of our circle and lines within R, i.e., we see intersections $C_n \cap R$ and $L \cap R$. Now clearly the segment $C_n \cap R$ gets closer to $L \cap R$ as n grows, and in the limit $n \to \infty$ they become identical. This is why we see the lines C_n converging to L. We see this convergence no matter how large (or small) the window

R is. Note that the Hausdorff distance between $C_n \cap R$ and $L \cap R$ indeed converges to zero: $\mathrm{dist}_{\mathrm{H}}(C_n \cap R, L \cap R) \to 0$ as $n \to \infty$.

To take care of any type of closed sets (both compact or noncompact), we change the definition of the classical Hausdorff distance between sets S_1 and S_2 as follows

$$\mathrm{dist}_{\mathrm{H}}(S_1, S_2; R) = \max\Big\{ \sup_{P \in S_1 \cap R} \mathrm{dist}(P, S_2), \sup_{Q \in S_2 \cap R} \mathrm{dist}(Q, S_1) \Big\} \tag{2.8}$$

if both S_1 and S_2, intersect the window R. Geometrically, it is the longest distance you have to travel from one set two another, provided you begin within R.

If only one set intersects R, say S_1, we can use the following expression:

$$\mathrm{dist}_{\mathrm{H}}(S_1, S_2; R) = \sup_{P \in S_1 \cap R} \mathrm{dist}(P, S_2). \tag{2.9}$$

or if nether set intersects R,

$$\mathrm{dist}_{\mathrm{H}}(S_1, S_2; R) = 0 \tag{2.10}$$

because we "see" two empty sets, which are not distinguishable.

W-convergence of sequences of sets (Main definition)

Now we use the Window-restricted Hausdorff distance to establish the convergence of sets. Our definition uses restricted window hence we will call the resulting notion "Window-convergence", or "W-convergence", for short.

Let $S_n \subset \mathbb{R}^2$ be some sets and $S \subset \mathbb{R}^2$ another set.

Definition 2.1. *The sequence S_n converges to S if for any finite window R we have*

$$\mathrm{dist}_{\mathrm{H}}(S_n, S; R) \to 0 \quad \text{as} \quad n \to \infty. \tag{2.11}$$

As it turns out that the sequence of lines L_n in the example indeed converges to the limit line L. like circles or ellipses, The W-convergence can be used equivalently as the convergence with respect to the Hausdorff distance.

The notion of convergence leads to constructions of topology and metric on the space of model objects. This is done in the next section.

2.3. Topology space of closed sets

In the last section we introduced the notion of convergence for a sequence of sets. Here we discuss various aspects of this new concept.

Uniqueness of a limit set Before we make any further discussion. Let us first consider such an example: let $L_n = \{y = x/n\}$ be a sequence of lines, converging to the x axis $L = \{y = 0\}$ and $L^{'}$ a subset of L consisting of all points whose coordinates are rational numbers. Then for any finite window R we have

$$\mathrm{dist}_{\mathrm{H}}(L_n, L^{'}; R) \to 0 \quad \text{as} \quad n \to \infty. \tag{2.12}$$

Thus, in the sense of W-convergence, the sequence L_n has two distinct limits: L and $L^{'}$. To avoid unnecessary complication, we will assume that all our sets $S \subset \mathbb{R}^2$ are closed. So we ignore any possible limits that is not closed. In this example L (a line) is closed, but the other limit set L' is not. In the problem of fitting curves, all model objects in - lines, circles, ellipses and other conics are closed sets.

Topology on the collection of objects

To induce the topological structure on the space of model object, one need to describe the collection of open sets. Let $\mathbb{X}$ denote the space whose elements are subsets $S \subset \mathbb{R}^2$ (and remember we agree to consider only closed sets S, so this will be assumed throughout). Having defined convergence of a sequence of sets S_n to a limit set S, we can define closed sets $Y \subset \mathbb{X}$ as follows: a set $Y \subset \mathbb{X}$ is closed if for any sequence of sets $S_n \in Y$ converging to a limit set S the limit set also belongs to Y, i.e., $S \in Y$. Now we have the collection of closed sets $Y \subset \mathbb{X}$. Then open sets $U \subset \mathbb{X}$ are those whose complements $\mathbb{X} \setminus U$ are closed, i.e., $U \subset \mathbb{X}$ is open if and only if $\mathbb{X} \setminus U$ is closed. One can easily check that the so defined open sets satisfy all axioms of a topological space.

W-distance between sets

It would be easier for us if we could quantify the W-convergence, i.e., if we could measure the distance between sets S_n and S in such a way that the W-convergence

$S_n \to S$ is equivalent to that the distance between S_n and S converges to zero. As we have seen, the Hausdorff distance from S_n to S would not do the job. Fortunately we can define a distance that will work. Let R_k denote a square window of size $2k \times 2k$, i.e.,

$$R_k = \{-k \le x \le k, \quad -k \le y \le k\}. \tag{2.13}$$

Now we define a *W-distance* (or a "Window-distance") between two sets $S_1, S_2 \subset \mathbb{R}^2$ as follows:

$$\mathrm{dist_W}(S_1, S_2) = \sum_{k=1}^{\infty} 2^{-k} \mathrm{dist_H}(S_1, S_2; R_k). \tag{2.14}$$

In this formula, we use a growing sequence of nested windows and the Hausdorff distances between S_1 and S_2 within those windows balanced by the factors 2^{-k}. In the formula (2.14), the first non-zero term corresponds to the smallest window R_k that intersects at least one of the two sets, S_1 or S_2. The sum in (2.14) is always finite. Indeed, let us suppose, for simplicity, that both S_1 and S_2 intersect each window R_k. Then (since the distance between any two points in R_k is at most $2\sqrt{2}k$) the above sum is bounded by $2\sqrt{2} \ \sum_{k=1}^{\infty} k2^{-k} < 6$.

Metrizable topology

Our W-distance (2.14) has the following property: given a sequence of sets S_n and a set S we have $\mathrm{dist_B}(S_n, S) \to 0$ as $n \to \infty$ if and only if $\mathrm{dist_H}(S_n, S; R_k) \to 0$ for each R_k (see proof in Appendix). Thus the sequence of sets S_n converges to a limit set S if and only if $\mathrm{dist_W}(S_n, S) \to 0$. This means that our topological space $\mathbb{X}$ of subsets of $\mathbb{R}^2$ can be completely described by W-distance (2.14). The fact that our space X *is* metrizable will be useful later.

We remark that our W-distance is constructed rather arbitrarily. First, it uses square windows, and we could have used rectangular or circular ones. Second, the windows are centered on the origin $(0, 0)$, while any other point would be just as good as a common center for our windows. Third, the factors 2^{-k} could be replaced with a^{-k} if we choose any other number $a > 1$, etc... etc... In fact the numerical value

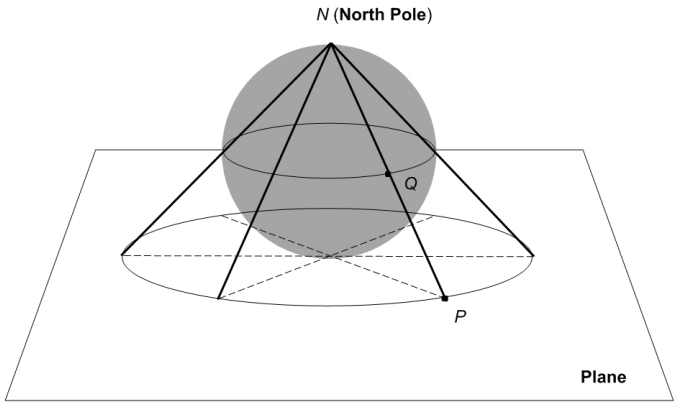


Figure 2.5: Riemann sphere

of our W-distance $\mathrm{dist}_{\mathrm{B}}(S_1, S_2)$ is pretty meaningless, the only important fact is that when S_n converges to S, then $\mathrm{dist}_{\mathrm{B}}(S_n, S) \to 0$, and vice versa.

Riemann sphere

Our concepts can be described differently if we map the plane $\mathbb{R}^2$ onto the Riemann sphere in the xyz space. Denote by

$$\mathbb{S} = \big\{(x, y, z)\colon x^2 + y^2 + (z - \tfrac{1}{2})^2 = \tfrac{1}{4}\big\} \tag{2.15}$$

the sphere in the xyz space (whose xy coordinate plane is our original plane $\mathbb{R}^2$) of radius $r = \frac{1}{2}$ centered on the point $(0, 0, \frac{1}{2})$. This is known as *Riemann sphere* (it is often used in complex analysis). It "rests" on the xy plane and its north pole $N = (0, 0, 1)$ is the highest point. Now every point $P = (x, y, 0) \in \mathbb{R}^2$ in the xy plane can be joined by a line with the north pole N of the sphere. This line PN intersects the sphere in a unique point $Q = \mathbb{S} \cap PN$ below the north pole. This defines a map $M\colon P \mapsto Q$ from the xy plane $\mathbb{R}^2$ onto the sphere $\mathbb{S}$. It can be visualized as the plane $\mathbb{R}^2$ "wrapped around" the sphere $\mathbb{S}$. It covers the entire sphere except the north pole N.

Every set $S \subset \mathbb{R}^2$ on the plane is thus mapped onto a set $S' = M(S) \subset \mathbb{S}$ on the sphere. If the set S is unbounded, then S' has the north pole N as a limit point,

and we need to add it to S' to make S' closed. Now given a sequence of closed sets $S_n \subset \mathbb{R}^2$ and a closed set $S \subset \mathbb{R}^2$ we get the corresponding closed sets $S'_n = M(S_n)$ and $S' = M(S)$ on the Riemann sphere $\mathbb{S}$. The convenience of this transformation is that we now can describe the W-convergence more easily than before: the sequence S_n converges (i.e., W-converges) to S if and only if the Hausdorff distance between their images on the sphere, S'_n and S', goes down to zero, i.e., $\mathrm{dist}_{\mathrm{H}}(S'_n, S') \to 0$, as $n \to \infty$. Thus if one uses the Riemann sphere as above, there is no need for our "window-restricted Hausdorff distances" or "W-distances", one can just refer to the regular Hausdorff distance on the Riemann sphere.

2.4. Continuity of the objective function

Our analysis of the problem of minimization of geometric distances from the given points to a model object is based on the continuity of the objective function, which allows us to use the classical extreme value theorem: *A continuous function on a nonempty compact space always attains its supremum and infimum.*

Objective function

The function to be minimized is the sum of squares of the distances from the given points to a model object:

$$\mathcal{F}(S) = \sum_{i=1}^{n} \big[\mathrm{dist}(P_i, S)\big]^2, \tag{2.16}$$

where $P_1, \ldots, P_n$ denote the given points and S a model object (from the given collection). The given points are fixed and they are not listed as arguments of $\mathcal{F}$. The given collection of model objects will be denoted by $\mathbb{M}$.

Model objects For the purpose of generality, we assume that model collection $\mathbb{M}$ contains some closed sets in $\mathbb{R}^2$. The reason for this requirement was explained in section 2.3. The collection $\mathbb{M}$ is then a subset of the topological space $\mathbb{X}$ of <u>all</u> closed sets in $\mathbb{R}^2$. The topology and metric in $\mathbb{X}$ were also introduced in section 2.3; now $\mathbb{M}$ automatically becomes a topological space and a metric space, too.

Redundancy principle

For any object $S' \subset S \in \mathbb{M}$ and a fixed point P we have

$$\operatorname{dist}(P, S) \le \operatorname{dist}(P, S'), \tag{2.17}$$

thus

$$\mathcal{F}(S) = \sum_{i=1}^{n} \big[\operatorname{dist}(P_i, S)\big]^2 \neq \mathcal{F}(S') \tag{2.18}$$

So for the purpose of minimizing $\mathcal{F}$, ignoring all proper sets may reduce the collection $\mathbb{M}$ somewhat. This reduction is not necessary, since there is no harm in considering any subset $S' \subset S$ of an object $S \in \mathbb{M}$ as a (smaller) object, too. If S' provides a best fit (i.e., minimizes the objective function $\mathcal{F}$), then so does S, because $\mathcal{F}(S) \le \mathcal{F}(S')$. Hence including S' into the collection $\mathbb{M}$ will not really be an extension of $\mathbb{M}$, its inclusion will not change the best fit.

Main Theorem (Continuity of the objective function)

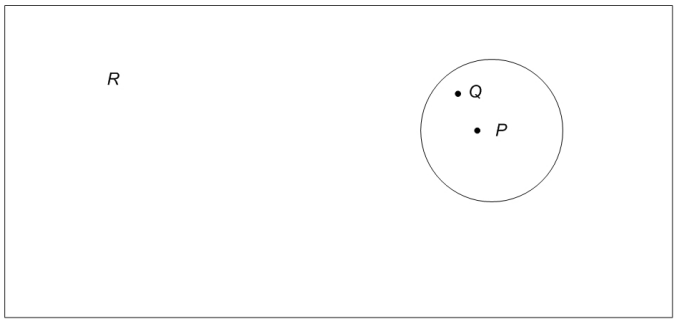
For any given points $P_1, \dots, P_n$ and any collection $\mathbb{M}$ of model objects (reminder: model objects are closed subsets of $\mathbb{R}^2$) the function $\mathcal{F}$ defined by (5.1) is continuous on $\mathbb{M}$. This means that if a sequence of objects $S_m \in \mathbb{M}$ converges (i.e., W-converges) to another object $S \in \mathbb{M}$, then $\mathcal{F}(S_m) \to \mathcal{F}(S)$.

Proof. Since $\mathcal{F}(S)$ is the sum of squares of distances $\operatorname{dist}(P_i, S)$ to individual points P_i, see (5.1), it is enough to verify that the distance $\operatorname{dist}(P, S)$ is a continuous function of S for every given point P.

Suppose we are given a point $P \in \mathbb{R}^2$ and a sequence of closed sets S_m W-converging to a set S. We denote by $Q \in S$ the point in S closest to P, i.e., such that

$$\operatorname{dist}(P, Q) = \operatorname{dist}(P, S), \tag{2.19}$$

see section 2.1. Denote by D the disk centered on P of radius $1 + \operatorname{dist}(P, Q)$; it contains Q. Let R be a window containing the disk D (windows were introduced in section (2.3)).



Since R contains Q, it intersects with S, i.e., $R \cap S \neq \emptyset$. This guarantees that $\mathrm{dist}_{\mathrm{H}}(S_m, S; R) \to 0$, according to section 2.3. Thus, there are points $Q_m \in S_m$ such that $Q_m \to Q$. Since $\mathrm{dist}(P, S_m) \leq \mathrm{dist}(P, Q_m)$, we conclude that the upper limit of the sequence $\mathrm{dist}(P, S_m)$ does not exceed $\mathrm{dist}(P, S)$, i.e.,

$$\limsup \mathrm{dist}(P, S_m) \leq \mathrm{dist}(P, S). \tag{2.20}$$

On the other hand, we will show that the lower limit of the sequence $\mathrm{dist}(P, S_m)$ cannot be smaller than $\mathrm{dist}(P, S)$, i.e.,

$$\liminf \mathrm{dist}(P, S_m) \geq \mathrm{dist}(P, S), \tag{2.21}$$

The estimates (2.20) and (2.21) together imply that $\mathrm{dist}(P, S_m) \to \mathrm{dist}(P, S)$, as desired, hence the distance function $\mathrm{dist}(P, S)$ will be continuous on $\mathbb{M}$. It remains to prove (2.21).

To prove (2.21), assume by way of contradiction that $\liminf \mathrm{dist}(P, S_m) < \mathrm{dist}(P, S)$. Then there is a subsequence S_{m_k} in our sequence of sets S_m such that

$$\lim_{k\to\infty} \mathrm{dist}(P, S_{m_k}) = \liminf \mathrm{dist}(P, S_m) < \mathrm{dist}(P, S). \tag{2.22}$$

Denote by $Q_m \in S_m$ the point in S_m closest to P, i.e., such that $\mathrm{dist}(P, Q_m) = \mathrm{dist}(P, S_m)$. Then we have

$$\lim_{k\to\infty} \mathrm{dist}(P, Q_{m_k}) = \lim_{k\to\infty} \mathrm{dist}(P, S_{m_k}) < \mathrm{dist}(P, S) = \mathrm{dist}(P, Q). \tag{2.23}$$

Since the points Q_{m_k} are closer to P than the point Q is, we have $Q_{m_k} \in D \subset R$. Recall that $\mathrm{dist}_{\mathrm{H}}(S_m, S; R) \to 0$, hence

$$\mathrm{dist}(Q_{m_k}, S) \to 0 \qquad \text{as} \quad k \to \infty. \tag{2.24}$$

Denote by $H_{m_k} \in S$ the point in S closest to Q_{m_k}, i.e., such that $\mathrm{dist}(Q_{m_k}, H_{m_k}) = \mathrm{dist}(Q_{m_k}, S)$. Now we have by triangle inequality

(2.25)

$$\mathrm{dist}(P, S) \le \mathrm{dist}(P, H_{m_k}) \le \mathrm{dist}(P, Q_{m_k}) + \mathrm{dist}(Q_{m_k}, H_{m_k}) = \mathrm{dist}(P, Q_{m_k}) + \mathrm{dist}(Q_{m_k}, S).$$

Now the limit of the first term on the right hand side of (2.25) is $< \mathrm{dist}(P, S)$ by (2.23), and the limit of the second term is zero by (2.24). This implies $\mathrm{dist}(P, S) < \mathrm{dist}(P, S)$, which is impossible. The contradiction proves (2.21). And the proof of (2.21) completes the proof of the theorem. □

We are now ready to proceed to the next section.

2.5. Closed collections of objects

Objective function

Recall that the function to be minimized is the sum of squares of the distances from the given points to a model object:

$$\mathcal{F}(S) = \sum_{i=1}^{n} \big[\mathrm{dist}(P_i, S)\big]^2, \tag{2.26}$$

where $P_1, \ldots, P_n$ denote the given points and S a model object from the given collection $\mathbb{M}$.

Best fitting object

Our goal is to choose $S_{\mathrm{best}} \in \mathbb{M}$ on which the function $\mathcal{F}$ takes its minimum value, i.e., such that

$$\mathcal{F}(S_{\mathrm{best}}) \le \mathcal{F}(S) \quad \text{for all} \quad S \in \mathbb{M}, \qquad \text{or} \qquad S_{\mathrm{best}} = \arg\min_{S \in \mathbb{M}} \mathcal{F}(S). \tag{2.27}$$

The model object S_{best} is called the best fit (or closest object) to the given points. Our fitting problem has a solution if S_{best} exists. Here we are preoccupied with the

existence of S_{best}. Does it always exist? If not, what issues can this cause? And how can we resolve them?

Infimum versus minimum

The function $\mathcal{F}$ defined by (5.1) cannot be negative, thus it always has a greatest lower bound :

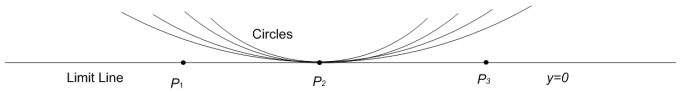
$$\mathcal{F}_0 = \inf_{S \in \mathbb{M}} \mathcal{F}(S). \tag{2.28}$$

So there always exists a sequence of objects S_n so that $\mathcal{F}(S_n)$ is arbitrarily close to $\mathcal{F}_0$. However the existence of the best fitting object S_{best} such that $\mathcal{F}(S_{\text{best}}) = \mathcal{F}_0$ is not guaranteed.

Practical issues In practical terms an algorithm that executes a certain iterative procedure produces a sequence of objects S_m (here m denotes the iteration number) such that $\mathcal{F}(S_m) < \mathcal{F}(S_{m-1})$, i.e., the quality of approximations improves with every step. If the procedure is successful, the value $\mathcal{F}(S_m)$ converges to the minimal possible value, $\mathcal{F}_0$, and the sequence of objects S_m converges (i.e., W-converges) to some limit object S_0. Then the continuity of the objective function (which we proved in section (2.4)) guarantees that $\mathcal{F}(S_0) = \mathcal{F}_0$, i.e., S_0 indeed provides the global minimum of the objective function, so it is the best fitting object: $S_0 = S_{\text{best}}$.

A problem arises if the limit object S_0 does *not* belong to the given collection $\mathbb{M}$, hence is not admissible. Then we end up with a sequence of objects S_m, each of which fits (approximates) the given points better than the previous one, but not as good as the next one. None of them would be the best fit, and in fact the best fit would not exist, so the fitting problem would have no solution.

Recall the example mentioned in the section 2.2. The sequence of circles S_m defined by $x^2 + (y - R)^2 = R^2$ will fit the points progressively better (tighter) as m grows, so that $\mathcal{F}(S_m) \to 0$ as $m \to \infty$. On the other hand, no circle can pass through three collinear points, hence no circle S satisfies $\mathcal{F}(S) = 0$. Thus the circle fitting problem has no solution in this case.



According to our section 2.2, the above sequence of circles S_m converges to the x axis, which is a line, so it is natural to declare that the line is the best fit. Though admittedly, in many practical applications one really needs to produce an estimate of the circle's center and radius. In that case a line would be of little help - it has no center or radius. But if we want to present the best fitting object here, it is clearly and undeniably the line $y = 0$.

Closed collections of objects

In order to guarantee the existence of the best fitting object in all cases, we need to include in our collection $\mathbb{M}$ all objects that can be obtained as limits of sequences of objects from $\mathbb{M}$. Such "limit objects" are *limit points* of $\mathbb{M}$, with respect to the defined topology. For example, The collection $\mathbb{M}_{\mathrm{L}}$ of all lines in $\mathbb{R}^2$ is closed, as a sequence of lines can only converge to a line. The collection $\mathbb{M}_{\mathrm{C}}$ of all circles in $\mathbb{R}^2$ is not closed, as a sequence of circles may converge to a circle or to a line. The closure of the collection of circles $\mathbb{M}_{\mathrm{C}}$ includes all circles and all lines, i.e.,

$$\bar{\mathbb{M}}_{\mathrm{C}} = \mathbb{M}_{\mathrm{C}} \cup \mathbb{M}_{\mathrm{L}}. \tag{2.29}$$

(Strictly speaking, a sequence of circles may also converge to a single point, see section 2.8, so singletons need to be included, too; this will be formally done later.) The collection of ellipses and that of hyperbolas will be investigated later in section 2.8.

Closedness is necessary

We see that the collection $\mathbb{M}$ of model objects must be closed if we want the best fitting object to exist in all cases. If $\mathbb{M}$ is not closed, we have to extend it by adding all its limit points and thus make it closed.

To justify the necessity of closedness more formally, suppose the collection $\mathbb{M}$ consists of curves of a certain type (lines, circles, conics, etc.) and it is not closed,

i.e., there is a sequence of objects $S_m \subset \mathbb{M}$ that converges to a closed set $S_0 \subset \mathbb{R}^2$ but S_0 is not included in $\mathbb{M}$. Let us place all the points $P_1, \ldots, P_n$ on the set S_0. If n is large enough and the points $P_1, \ldots, P_n$ are distinct, then there is at most one object of the given type that interpolates all the selected points $P_1, \ldots, P_n$ (for example, there is a unique line interpolating two distinct points, at most one circle interpolating three distinct points, etc.). Since the object S_0 is not included in $\mathbb{M}$, there is no object $S \in \mathbb{M}$ that interpolates our points, i.e., $\mathcal{F}(S) \neq 0$ for any $S \in \mathbb{M}$. On the other hand, since $S_m \to S$, we have $\mathcal{F}(S_m) \to 0$ as $m \to \infty$. We see that the function $\mathcal{F}$ fails to take its minimum value (zero), hence the best fitting object (for which $\mathcal{F}(S) = 0$) does not exist in $\mathbb{M}$. So a single object S_0 that is not included in $\mathbb{M}$ may cause the failure of the function $\mathcal{F}$ to take its minimum.

Closedness is sufficient

To ensure the existence of best fit in the collection $\mathbb{M}$ of model objects $\mathbb{M}$, the closedness is sufficient - the best fitting object always exists whenever $\mathbb{M}$ is closed. Due to the importance of this fact, we will prove it in the next section.

2.6. Existence of the best fit

The purpose of this section is to prove that if the collection M of model objects is closed, then the best fitting object exists (for any set of given points $P_1, \ldots P_n$). In other words, the objective function $\mathcal{F}$ always attains its global minimum.

The key ingredients of our proof will be the continuity of the objective function and the compactness of a restricted domain of that function. The following general fact will be used:

A continuous real-valued function on a compact set always takes its maximum value and its minimum value on that set.

Non-compactness of $\mathbb{M}$

In metric spaces like our $\mathbb{M}$, a subset $\mathbb{M}_0 \subset \mathbb{M}$ is compact if every sequence of its elements $S_m \in \mathbb{M}_0$ has a subsequence S_{m_k} that converges to another element $S \in \mathbb{M}_0$, i.e., $S_{m_k} \to S$ as $k \to \infty$.

We know that our objective function $\mathcal{F}$ is continuous; see section 2.4. Its domain $\mathbb{M}$ is now assumed to be closed. If it *was* compact, the above general fact would guarantee that $\mathcal{F}$ had a global minimum, as desired.

But is $\mathbb{M}$ compact? We can check this by referring to the above theorem again. If it *was* compact, the function $\mathcal{F}$ would take both a minimum, and a maximum. And this is impossible since the model objects can move arbitrarily far from the given points, thus $\mathcal{F}(S) \to +\infty$.

Necessity of a restricted collection

The reason why $\mathbb{M}$ fails to be compact is that it is "too large". It contains model sets $S \in \mathbb{M}$ that are too far from the given points. It is exactly those objects which prevent $\mathbb{M}$ from being compact: any sequence of model objects located farther and farther away from the given points would "escape to infinity", rather than converge to any object S. So we need to find a smaller (restricted) subcollection $\mathbb{M}_0 \subset \mathbb{M}$ which will be compact and then we will apply the above general fact.

Construction of a restricted collection

For a set of given points $P_1, \dots, P_n$, find an $r > \max_{i=1,\dots,n} \operatorname{dist}(P_i, (0,0))$. Besides, let us assume that r is large enough so that the disk of radius r $D_r = \{x^2 + y^2 \le r^2\}$ centered on the origin $(0,0)$ intersect at least one object $S_0 \in \mathbb{M}$. The distances from the given points to S_0 cannot exceed the diameter of D_r, which is $2r$, hence $\mathcal{F}(S_0) \le (2r)^2 n$.

Now we define our subcollection $\mathbb{M}_0 \subset \mathbb{M}$: it consists of all model objects $S \in \mathbb{M}$ that intersect the larger disk D_{3r} of radius $3r$. Objects that lie entirely outside D_{3r} are not included in $\mathbb{M}_0$. Note that the subcollection $\mathbb{M}_0$ contains at least one object: it contains S_0 mentioned above, because S_0 intersects the smaller disk D_r. Hence $\mathbb{M}_0$ is not empty.

Restriction to $\mathbb{M}_0$

Recall that all our given points $P_1, \ldots, P_n$ lie in D_r. They are separated from the region outside the larger disk D_{3r} by the ring $D_{3r} \setminus D_r$, which is $2r$ wide. Thus the distances from the given points to any object S which was not included in $\mathbb{M}_0$ are greater than $2r$, hence for such objects we have $\mathcal{F}(S) > (2r)^2 n$. Hence objects not included in $\mathbb{M}_0$ cannot fit our points better than S_0 does. Therefore they can be ignored in the process of minimization of $\mathcal{F}$. More precisely, if we find the best fitting object S_{best} within the subcollection $\mathbb{M}_0$, then for any other object $S \in \mathbb{M} \setminus \mathbb{M}_0$ we will have

$$\mathcal{F}(S) > (2r)^2 n \geq \mathcal{F}(S_0) \geq \mathcal{F}(S_{\text{best}}), \tag{2.30}$$

which shows that S_{best} will be also the best fitting object in the entire collection $\mathbb{M}$.

Compactness of $\mathbb{M}_0$

It remains to check that the subcollection $\mathbb{M}_0$ is compact. Recall that the non-compactness of the original collection $\mathbb{M}$ was caused by sequences of objects $S_m \in \mathbb{M}$ located progressively farther away from the given points (sequences of objects that "escape to infinity"). In the subcollection $\mathbb{M}_0$ such sequences are impossible: all objects $S \in \mathbb{M}_0$ are required to intersect the disk D_{3r}, so they are all at a distance $< 6r$ from the given points (in fact, the distance is $< 4r$ because the given points are all in the smaller disk D_r).

Now since the subcollection $\mathbb{M}_0$ is compact and the objective function $\mathcal{F}$ is continuous, it takes a minimum value by the above theorem, hence the best fitting object exists.

Formal proof of compactness of $\mathbb{M}_0$ The above argument is rather informal. We need to verify that every sequence of objects $S_m \in \mathbb{M}_0$ has a subsequence converging to another object $S^* \in \mathbb{M}_0$.

We will use the following general fact: **In a compact metric space any sequence of compact subsets has a convergent subsequence with respect to the Hausdorff distance** .

Main Theorem (Existence of the best fit) Suppose the given collection $\mathbb{M}$ of model objects is closed (see below). Then for any given points $P_1, \ldots, P_n$ there exists the best fitting object $S_{\text{best}} \in \mathbb{M}$. This means that

$$\mathcal{F}(S_{\text{best}}) \le \mathcal{F}(S) \quad \text{for all} \quad S \in \mathbb{M}, \qquad \text{or} \qquad S_{\text{best}} = \arg\min_{S \in \mathbb{M}} \mathcal{F}(S), \tag{2.31}$$

i.e., the objective function $\mathcal{F}$ attains its global minimum on $\mathbb{M}$. Recall: A collection $\mathbb{M}$ of model objects is closed if for any sequence of objects $S_m \in \mathbb{M}$ that W-converges to an object S, the limit object S also belongs to $\mathbb{M}$.

PROOF. Now let $j = [3r] + 1$ be the smallest integer greater than 3r. Recall that all the objects in $\mathbb{M}_0$ are required to intersect D_{3r}, thus they all intersect D_j as well. The sets $S_m \cap D_j$ are compact. By the above general fact, there is a subsequence $S_k^{(j)}$ in the sequence S_m and a compact subset $S_j^* \subset D_j$ such that $\text{dist}_{\text{H}}(S_k^{(j)} \cap D_j, S_j^*) \to 0$ as $k \to \infty$. Next, from the subsequence $S_k^{(j)}$ we extract a subsequence, call it $S_k^{(j+1)}$ that converges in the larger disk D_{j+1}, i.e., such that $\text{dist}_{\text{H}}(S_k^{(j+1)} \cap D_{j+1}, S_{j+1}^*) \to 0$ as $k \to \infty$ for some compact subset $S_{j+1}^* \subset D_{j+1}$. Since $\text{dist}_{\text{H}}(S_k^{(j+1)} \cap D_j, S_j^*) \to 0$, we see that $S_{j+1}^* \cap D_j = S_j^*$, i.e., the limit sets S_j^* and S_{j+1}^* "agree" within D_j.

Then we continue this procedure inductively for the progressively larger disks $D_{j+2}, D_{j+3}, \ldots$. In the end we use standard Cantor's diagonal argument [**19**] to construct a single subsequence S_{m_k} such that for every $i \ge j$ we have $\text{dist}_{\text{H}}(S_{m_k} \cap D_i, S_i^*) \to 0$ as $k \to \infty$, and the limiting subsets $S_i^* \subset D_i$ agree in the sense $S_{i+1}^* \cap D_i = S_i^*$ for every $i \ge j$. Then it follows that the sequence of objects S_{m_k} converges (i.e., W-converges) to the closed set $S^* = \cup_{i \ge j} S_i^*$. The limit set S^* must belong to our collection $\mathbb{M}$ because that collection was assumed to be closed. Lastly, since S^* intersects the disk D_{3r}, it also belongs to the subcollection $\mathbb{M}_0$. Our formal proof is now complete. □

A different approach to the existence problem is given in the next Section.

2.7. Megaspace

Introduction

Our conclusions can be illustrated by an interesting construction in a multidimensional space ("megaspace"). It was first used by Malinvaud [**24**] and then by Chernov (Section 1.5 and Section 3.4 in [**12**]).

Objective function (reminder)

Recall that given n data points $P_1 = (x_1, y_1)$, ..., $P_n = (x_n, y_n)$ and a model object S (a closed subset of $\mathbb{R}^2$), the objective function $\mathcal{F}(S)$ is defined by

$$\mathcal{F}(S) = \sum_{i=1}^{n} \big[\mathrm{dist}(P_i, S)\big]^2$$

and for the distance $\mathrm{dist}(P_i, S)$ we can write

$$\big[\mathrm{dist}(P_i, S)\big]^2 = \min_{(x_i', y_i') \in S} \big\{(x_i - x_i')^2 + (y_i - y_i')^2\big\}.$$

Thus we can express the objective function $\mathcal{F}(S)$ as follows:

$$\mathcal{F}(S) = \min\left\{\sum_{i=1}^{n}\big[(x_i - x_i')^2 + (y_i - y_i')^2\big]; \qquad (x_i', y_i') \in S \quad \forall i\right\}. \tag{2.32}$$

Megapoints and megasets Now let us represent the n data points $P_1 = (x_1, y_1)$, ..., $P_n = (x_n, y_n)$ by <u>one</u> point ("megapoint") $\mathcal{P}$ in the $2n$-dimensional space $\mathbb{R}^{2n}$ with coordinates $x_1, y_1, \ldots, x_n, y_n$.

For the given object S, let us also define a multidimensional set ("megaset") $\mathfrak{M}_S \subset \mathbb{R}^{2n}$ as follows:

$$\mathcal{P}' = (x_1', y_1', \ldots, x_n', y_n') \in \mathfrak{M}_S \quad \Longleftrightarrow \quad (x_i', y_i') \in S \qquad \forall i.$$

Note that $\sum_{i=1}^{n}\big[(x_i - x_i')^2 + (y_i - y_i')^2\big]$ in (2.32) is the square of the distance from $\mathcal{P}$ to $\mathcal{P}' \in \mathfrak{M}_S$, in the Euclidean metric in the "megaspace" $\mathbb{R}^{2n}$. Therefore

$$\mathcal{F}(S) = \min_{\mathcal{P}' \in \mathfrak{M}_S} \big[\mathrm{dist}(\mathcal{P}, \mathcal{P}')\big]^2 = \big[\mathrm{dist}(\mathcal{P}, \mathfrak{M}_S)\big]^2. \tag{2.33}$$

Next given a collection $\mathbb{M}$ of model objects we define a large "megaset" $\mathfrak{M}(\mathbb{M}) \subset \mathbb{R}^{2n}$ as follows: $\mathfrak{M}(\mathbb{M}) = \cup_{S \in \mathbb{M}} \mathfrak{M}_S$. Alternatively, it can be defined as

$$(x_1^{'}, y_1^{'}, \ldots, x_n^{'}, y_n^{'}) \in \mathfrak{M}(\mathbb{M}) \iff \exists S \in \mathbb{M} \colon (x_i^{'}, y_i^{'}) \in S \qquad \forall i.$$

We will describe the "megasets" $\mathfrak{M}(\mathbb{M})$ for several commonly used model collections $\mathbb{M}$ (of lines, circles, and ellipses) in section 2.10.

Projecting megapoint onto megaset The best fitting object S_{best} minimizes the function $\mathcal{F}(S)$. Thus, due to (2.33), S_{best} minimizes the distance from the "megapoint" $\mathcal{P}$ representing the data set $P_1, \ldots, P_n$ to the "megaset" $\mathfrak{M}(\mathbb{M})$ representing the collection $\mathbb{M}$.

Thus, the problem of finding the best fitting object S_{best} reduces to the problem of finding the "megapoint" $\mathcal{P}' \in \mathfrak{M}(\mathbb{M})$ that is the closest to the given "megapoint" $\mathcal{P}$ (representing the given n points). In geometric terms, we need to project the "megapoint" $\mathcal{P}$ onto the "megaset" $\mathfrak{M}(\mathbb{M})$, and the footpoint $\mathcal{P}'$ of the projection would give us the best fitting object S_{best}.

Closedness Conclusion: the best fitting object $S_{\text{best}} \in \mathbb{M}$ exists if and only if there exists a "megapoint" $\mathcal{P}' \in \mathfrak{M}(\mathbb{M})$ that is the closest to the given "megapoint" $\mathcal{P}$.

It is a simple fact in geometry (already mentioned in Section (2.1)) that given a set $D \subset \mathbb{R}^d$ in a Euclidean space, the closest point $X' \in D$ to a given point $X \in \mathbb{R}^d$ exists for any X if and only if the set D is *closed* (in topological sense).

Thus the existence of the best fitting object S_{best} requires the "megaset" $\mathfrak{M}(\mathbb{M})$ to be topologically closed. Again we see that the property of closedness is necessary and sufficient for the fitting problem to have a solution.

The closedness of the model collection $\mathbb{M}$ guarantees the closedness of the megaset $\mathfrak{M}(\mathbb{M})$:

THEOREM 2.1. *If the model collection* $\mathbb{M}$ *is closed (in the sense of the W-convergence, as defined in Section 2.5), then the megaset* $\mathfrak{M}(\mathbb{M})$ *is closed in the natural topology of* $\mathbb{R}^{2n}$.

The significance of this theorem is clear as it provides an alternative proof of the existence of the best fitting object, provided the model collection $\mathbb{M}$ is closed. The proof of the above theorem is given next.

PROOF. Suppose a sequence of megapoints

$$\mathcal{P}^{(k)} = (x_1^{(k)}, y_1^{(k)}, \ldots, x_n^{(k)}, y_n^{(k)}),$$

all belonging to the megaset $\mathfrak{M}(\mathbb{M})$, converges, as $k \to \infty$, to a megapoint

$$\mathcal{P}^{(\infty)} = (x_1^{(\infty)}, y_1^{(\infty)}, \ldots, x_n^{(\infty)}, y_n^{(\infty)})$$

in the usual topology of $\mathbb{R}^{2n}$. This implies that for every $i = 1, \ldots, n$ the point $(x_i^{(k)}, y_i^{(k)})$ converges, as $k \to \infty$, to the point $(x_i^{(\infty)}, y_i^{(\infty)})$. We need to show that $\mathcal{P}^{(\infty)} \in \mathfrak{M}(\mathbb{M})$.

Now for every k there exists a model object $S^{(k)} \in \mathbb{M}$ that contains the n points $(x_1^{(k)}, y_1^{(k)}), \ldots, (x_n^{(k)}, y_n^{(k)})$. The sequence $\{S^{(k)}\}$ contains a convergent subsequence S^{k_r}, i.e., such that $S^{k_r} \to S$ as $r \to \infty$ for some closed set $S \subset \mathbb{R}^2$. The reason for the existence of a convergent subsequence is the same as in the proof of the compactness of $\mathbb{M}_0$ in the end of Section 2.6.

As we assumed that the collection $\mathbb{M}$ is closed, it follows that $\mathbb{M}$ contains the limit object S, i.e., $S \in \mathbb{M}$. It is intuitively clear (and can be checked by a routine calculus-type argument) that S contains all the limit points $(x_1^{(\infty)}, y_1^{(\infty)}), \ldots, (x_n^{(\infty)}, y_n^{(\infty)})$. Therefore the limit megapoint $\mathcal{P}^{(\infty)}$ belongs to the megaset $\mathfrak{M}(\mathbb{M})$. This proves that the megaset is closed. □

Two versions of closedness We defined the notion of "closed collections of model objects" in the section 2.5. We proved that the closedness of the collection of model objects is necessary and sufficient for the existence of the best fitting object.

Here we introduced another notion of closedness: given a collection of model objects $\mathbb{M}$ we constructed its megaset $\mathfrak{M}(\mathbb{M}) \subset \mathbb{R}^{2n}$, and then used the closedness of the megaset (with respect to the natural topology in $\mathbb{R}^{2n}$) to justify the existence of the best fitting object.

A natural question is: Are these two types of closedness equivalent? We have shown that the closedness of $\mathbb{M}$ implies that of $\mathfrak{M}(\mathbb{M})$. Curiously, the converse is not true. It may happen that $\mathfrak{M}(\mathbb{M})$ is closed, but $\mathbb{M}$ is not, see examples below. So these two versions of closedness are not equivalent.

We will argue, however, that those examples are in a way exceptional, and under certain reasonable conditions the two versions of closedness are indeed equivalent.

Examples of exceptional cases First we point out some exceptions. For example, let $\mathbb{M}$ be the collection of circles. It is not closed as we learned in section 2.8. Suppose $n = 2$. Since every two points belong to some circle, we have $\mathfrak{M}(\mathbb{M}) = \mathbb{R}^4$, which is a closed set. Thus it is possible that $\mathbb{M}$ is not closed, but $\mathfrak{M}(\mathbb{M})$ is closed.

As a more sophisticated example, let $\mathbb{M}$ be the collection of the following objects: all circles, all lines (with the exception of the x-axis), all singletons, all three-point sets in the x-axis, and all two-point sets in the x-axis. We note that all sets $\{(x_1, 0),\ (x_2, 0),\ (x_3, 0)\}$, where x_1, x_2, x_3 are arbitrary numbers (distinct or equal), are included in $\mathbb{M}$. We also note that this collection is not closed, because a sequence of lines may converge to the x-axis, which does not belong to $\mathbb{M}$ (as it was specifically excluded).

Now let $n = 3$. It is easy to check that "any" three points belong to one of the objects in our collection $\mathbb{M}$. Therefore, we have $\mathfrak{M}(\mathbb{M}) = \mathbb{R}^6$, which is a closed set. Thus again, $\mathbb{M}$ is not closed, but $\mathfrak{M}(\mathbb{M})$ is closed.

The above examples are rather artificial, as it is unnatural to fit circles to $n = 2$ or $n = 3$ points. This is rather an interpolation problem, not a fitting problem. This observation will help us to find a general approach to the issue.

Equivalence of two versions of closedness Suppose that our collection of model objects $\mathbb{M}$ has the following property: there exists $n_0 \geq 1$ such that any two distinct objects can intersect in at most n_0 points. This means that if $S_1, S_2 \in \mathbb{M}$ are two distinct objects, i.e., $S_1 \neq S_2$, then their intersection $S_1 \cap S_2$ is a finite set consisting of at most n_0 points.

For lines $n_0 = 1$, for circles $n_0 = 2$, for ellipses and other conics $n_0 = 4$ [**?**]. If such n_0 exists, then for any $n_0 + 1$ distinct points there can be at most one object containing all of those $n_0 + 1$ points.

The number n_0+1 can be described as follows: any object $S \in \mathbb{M}$ can be identified (i.e., uniquely determined) by any n_0+1 distinct points in it. Any line can be identified by two distinct points on it, any circle - by three distinct points on it, etc.

Now the equivalence of the above two versions of closedness holds only if $n > n_0+1$. For example, if our model collection consists of lines, then the equivalence holds for $n > 2$. If our collection consists of circles, then $n > 3$, for ellipses and other conics the requirement is $n > 5$.

We also need to assume a natural condition: if a sequence of objects $S_m \in \mathbb{M}$ converges (in our sense, i.e., W-converges) to a closed set S^*, i.e., $S_m \to S^*$, and S^* is a subset of an object $S \in \mathbb{M}$, i.e., $S^* \subset S$, then $S^* \in \mathbb{M}$. This assumption agrees with the Redundancy Principle, see section 2.4, as any subset of a legitimate object $S \in \mathbb{M}$ can be regarded as an object, too, for fitting purposes.

THEOREM 2.2. *(Equivalence of two versions of closedness) Let $\mathbb{M}$ be a collection of model objects. Suppose any two distinct objects $S_1, S_2 \in \mathbb{M}$ can intersect each other in at most n_0 points. Then $\mathbb{M}$ is closed (in the sense of section 2.5) if and only if for any $n > n_0 + 1$ the corresponding megaset $\mathfrak{M}(\mathbb{M}) \subset \mathbb{R}^{2n}$ is closed with respect to the natural topology in $\mathbb{R}^{2n}$.*

PROOF. Part (i): If $\mathbb{M}$ is closed, then $\mathfrak{M}(\mathbb{M})$ is closed.

By way of contradiction, let us assume that $\mathfrak{M}(\mathbb{M})$ is not closed (in a natural sense). This indicates that $\mathfrak{M}(\mathbb{M})$ does not contain one of its limit point $\mathcal{P}_0$. Then

there exists a sequence of megapoints

$$\mathcal{P}_k = (x_{k1}, y_{k1}, \ldots, x_{kn}, y_{kn}) \in \mathfrak{M}(\mathbb{M}) \qquad (x_{ki}, y_{ki}) \in S_k \in \mathbb{M} \qquad \forall i.$$

converging to $\mathcal{P}_0 = (x_1, y_1, \ldots, x_n, y_n)$. In addition, $\left[\text{dist}(\mathcal{P}_k, (0, 0, \ldots, 0, 0))\right]^2 < D$ $(D \in R^+)$ for all k. Let $D_M = \{x^2 + y^2 \leqslant D\}$. So D_M intersects with all S_k and by a similar argument as section 2.6, the subset of $\mathbb{M}$ containing every S_k is compact. Let's just assume $\{S_k\}$ converges to a finite limit S_0. Because of closedness of $\mathbb{M}$, $S_0 \in \mathbb{M}$. Since $\mathcal{P}_0$ does not belong to $\mathfrak{M}(\mathbb{M})$, $\text{dist}(S_0, (x_i, y_i)) > 0$ for some i. For a finite window $R = \{-\sqrt{D} \leq x \leq \sqrt{D}, \ \ -\sqrt{D} \leq y \leq \sqrt{D}\}$ which contains every (x_{ki}, y_{ki}). Therefore

$$\text{dist}((x_{ki}, y_{ki}), (x_i, y_i)) > \text{dist}(S_0, (x_i, y_i)) - \varepsilon$$

which contradicts the fact that $\mathcal{P}_k$ converges to $\mathcal{P}_0$. The contradiction proves (i).

Part (ii): if $\mathfrak{M}(\mathbb{M})$ is closed, then $\mathbb{M}$ is closed.

Assume that one of the limit points S_0 (a nonempty closed set) of $\mathbb{M}$ does not belong to $\mathbb{M}$. In precise terms there exists a sequence $\{S_k\}$ $(k = 1, 2, \ldots)$ in $\mathbb{M}$ converging to a limit $S_0 \notin \mathbb{M}$ in the space of all closed sets in $\mathbb{R}^2$). First, if S_0 contains more than n distinct points, pick n distinct points $(x_i, y_i) \in S_0$ $(i = 1, \ldots, n)$. Let $R = \{-A \leq x \leq A, \ \ -B \leq y \leq B\}$ be a rectangle containing all data points (x_i, y_i) $(i = 1, \ldots, n)$. For any $\varepsilon > 0$, $\text{dist}_{\text{H}}(S_k, S_0; R) < \varepsilon$ for large enough k. Let (x_{ki}, y_{ki}) be the closest point in S_k to (x_i, y_i). Then

$$\text{dist}((x_i, y_i), S_k) = \text{dist}((x_i, y_i), (x_{ki}, y_{ki})) \leqslant \text{dist}_{\text{H}}(S_k, S_0; R) < \varepsilon$$

It follows that the sequence of megapoints

$$\mathcal{P}_k = (x_{k1}, y_{k1}, \ldots, x_{kn}, y_{kn}) \to \mathcal{P} = (x_1, y_1, \ldots, x_n, y_n) \quad \text{ as } \ k \to \infty.$$

Since $\mathfrak{M}(\mathbb{M})$ is closed and $\mathcal{P}_k \in \mathfrak{M}(\mathbb{M})$, $\mathcal{P} \in \mathfrak{M}(\mathbb{M})$. Then there exists a model object $S_0^{'} \in \mathbb{M}$ containing (x_i, y_i) $(i = 1, \ldots, n)$. Remember that we assumed a natural condition: if a sequence of objects $S_m \in \mathbb{M}$ converges to a closed set S^*, i.e.,

$S_m \to S^*$ (in the sense of section 2.5), and S^* is a subset of an object $S \in \mathbb{M}$, i.e., $S^* \subset S$, then $S^* \in \mathbb{M}$. So S_0 can not be the subset of $S_0^{'}$ and there exists a point $(x_n^{'}, y_n^{'})$ in S_0 but not belonging to $S_0^{'}$. Let us consider another set of n distinct points (x_i, y_i) $(i = 1, \ldots, n-1)$ and $(x_n^{'}, y_n^{'})$. It is easy to see that there exists an $S_0^* \in \mathbb{M}$ containing (x_i, y_i) $(i = 1, \ldots, n-1)$ and $(x_n^{'}, y_n^{'})$. So

$$S_0^{'} \cap S_0^* = \{(x_i, y_i), i = 1, \ldots, n-1\} \tag{2.34}$$

But recall that $n - 1 > n_0$. Thus (2.34) contradicts our assumption that any two distinct objects in $\mathbb{M}$ can intersect each other in at most n_0 points.

Next, suppose S_0 contains no more than n points (x_i, y_i) $(i = 1, \ldots, l \leqslant n)$. Let us define a megapoint $\mathcal{P}$ by those l points and another $n-l$ identical points at (x_l, y_l). So

$$P = (x_1, y_1, \ldots, x_l, y_l, \ldots, x_l, y_l)$$

where (x_l, y_l) is repeated $n-l+1$ times. By similar argument as above, one can find a sequence of megapoints

$$\mathcal{P}_k = (x_{k1}, y_{k1}, \ldots, x_{kn}, y_{kn}) \to \mathcal{P} = (x_1, y_1, \ldots, x_l, y_l, \ldots, x_l, y_l) \in \mathfrak{M}(\mathbb{M})$$

Then there exists a model object $S_0^{'} \in \mathbb{M}$ containing (x_i, y_i) $(i = 1, \ldots, l)$. This shows that $S_0 \subset S_0^{'}$ and $S_0 \in \mathbb{M}$, which contradicts that S_0 does not belong to $\mathbb{M}$. Therefore, $\mathbb{M}$ must contain all its limit points and it is closed. □

2.8. Model objects in 2D

In this section we will discuss the issue of existence for several popular models widely used in practical applications: ellipses, hyperbolas, parabolas. The same type of issues concerning circle fitting has been resolved and some exceptional cases can be constructed to disclose a fact that there are some data sets for which the best fitting circle that minimizes the sum of squares of orthogonal distances does not exists. Recall an example of n $(n \geq 3)$ collinear data points. One can approximate those

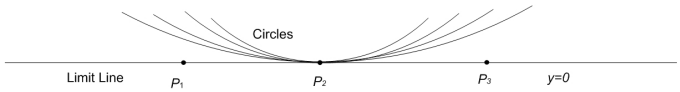


Figure 2.6: a sequence of circles converges to a line

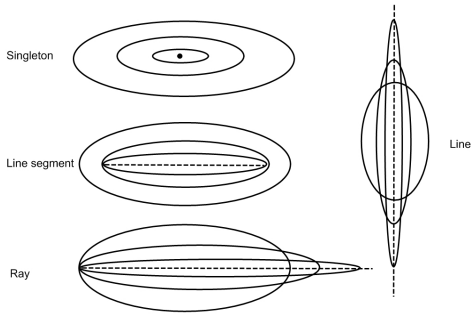


Figure 2.7: a sequence of ellipses converges to a singleton, line, ray or line segment

points with a large circular arc and make the sum of squares arbitrarily small but no the best fit will never be achieved (see Figure 2.6).

So no particular circle will provide the global minimum for $\mathcal{F}$. This suggests that nonexistence of the best fitting may arise when one tries to fit ellipse, parabola and hyperbola as well. We will show all possible limiting object approached by our fitting models and provide the sufficient condition for existence of the best fit.

Ellipses Fitting ellipses to data points is a common task in computer vision. Given data points $P_1, \ldots, P_n$, the best fitting ellipse E_{best} minimizes the sum of squares of the distances from $P_1, \ldots, P_n$ to the ellipse. The model collection $\mathbb{M}_{\text{E}}$ consists of all ellipses $E \subset \mathbb{R}^2$ have several different types of limiting objects other than ellipse and the collection $\mathbb{M}_{\text{E}}$ of ellipses is not closed.

First, the limit object may be an ellipse $E_0 \subset \mathbb{R}^2$. Since circles are also ellipses, lines and singletons as limit points in the model collection of all circles should be also included as the limit point in the closure of model collection of ellipses. Next, when a sequence of ellipse with minor axis shrinking to 0, the limit object will be a line, line segment or ray. They can be regarded as the degenerate ellipses with minor axis 0.

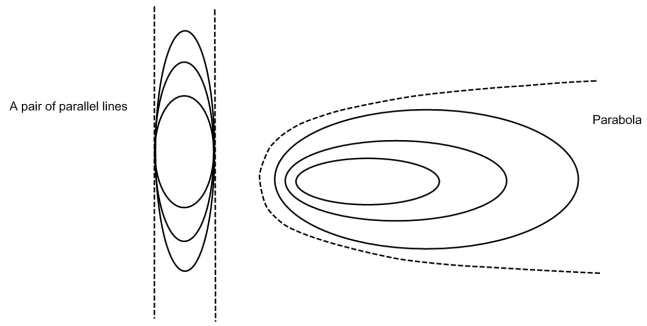


Figure 2.8: a sequence of ellipses converges to a pair of parallel lines or parabola

Besides, the ellipse can also converge to a pair of parallel lines if the major axis extend to infinity and minor axis stay constant or parabolas if both semi-axis extend to infinity.

Since both singletons and line segments, as well as rays are part of lines, they can be ignored based on the redundancy principle; see section 2.4. But lines, pairs of parallel lines, and parabolas cannot be brushed off so easily.

Ellipses (extended)

By adding all the limit points (lines, pairs of parallel lines, parabola) into the collection $\mathbb{M}_{\mathrm{E}}$, we extended the original set into its closure denoted by $\bar{\mathbb{M}}_{\mathrm{E}}$:

$$\bar{\mathbb{M}}_{\mathrm{E}} = \mathbb{M}_{\mathrm{E}} \cup \mathbb{M}_{\mathrm{L}} \cup \mathbb{M}_{\parallel} \cup \mathbb{M}_{\cup}, \tag{2.35}$$

where $\mathbb{M}_{\parallel}$ denotes the collection of pairs of parallel lines, and $\mathbb{M}_{\cup}$ the collection of parabolas.

By our main theorem (2.6), for any data points $P_1, \ldots, P_n$ there exists a best fitting object $S_{\mathrm{best}} \in \bar{\mathbb{M}}_{\mathrm{E}}$ where the objective function (5.1) achieves its minimum. But keep in mind, though, that the best fitting object may be a line, or a pair of parallel lines, or a parabola, rather than an ellipse. We summarize this result as follows:

THEOREM 2.3. *Let B be a given compact set (Euclidean space) $\mathbb{B}$ containing all the data points. Then the 'enlarged' space Ω of ellipses, parabolas, lines(including rays*

and line segments, singletons), and pairs of parallel lines intersecting B is compact with respect to the topology defined on the Ω (see section 2.3).

See section A.2 in appendix for the proof of the theorem.

Extension for ellipses: history

The ellipse fitting problem has been around since the 1970s. The need to deal with limiting cases was first notice by Bookstein [**8**], who wrote: " The fitting of a parabola is a limiting case, exactly transitional between ellipse and hyperbola. As the center of ellipse moves off toward infinity while its major axis and the curvature of one end are held constant... " A first theoretical analysis on the non-existence of the best ellipse was done by Nievergelt in [**26**] who traced it to the non-compactness of the underlying model space and concluded that parabolas needed to be included in the model space to guarantee the existence of the best fit.

Quadratic curves Now we deal with a model collection which consists of all quadratic curves, by which we mean all ellipses, parabolas and hyperbolas. Apparently, in order to ensure the best fitting curve, one needs to include all types of model objects in $\bar{\mathbb{M}}_{\mathrm{E}}$ into the extended collection. In fact, we only need to investigate the limit objects of hyperbolas. The hyperbola may converge to a pair of intersecting lines or two opposite half-lines (rays)(see illustration in 2.10). Since half-line and two opposite half-lines are a part (subset) of a full line, they can be ignored based on the redundancy principle; see section 2.4. Besides, the hyperbola can also transform into two parallel lines, single lines, rays or parabola, which are also limit objects included in $\bar{\mathbb{M}}_{\mathrm{E}}$

Quadratic curves (extended) We denote that extended collection by $\bar{\mathbb{M}}_{\mathrm{Q}}$:

$$\bar{\mathbb{M}}_{\mathrm{Q}} = \mathbb{M}_{\mathrm{Q}} \cup \mathbb{M}_{\mathrm{L}} \cup \mathbb{M}_{\parallel} \cup \mathbb{M}_{\times}, \tag{2.36}$$

where $\mathbb{M}_{\parallel}$ was introduced in (2.35) and $\mathbb{M}_{\times}$ denotes the collection of pairs of intersecting lines.

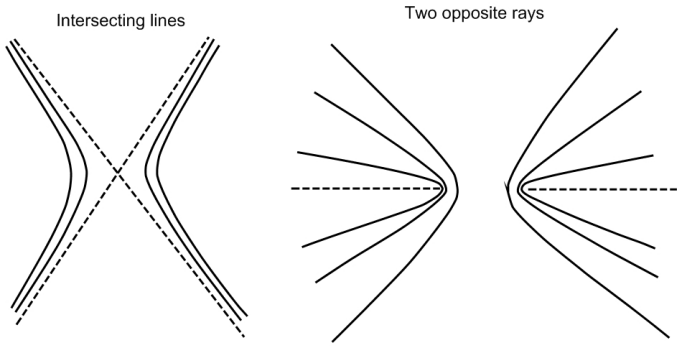


Figure 2.9: a sequence of hyperbola converges to a pair of intersecting lines or opposite rays

Now by our main theorem 2.6,the (extended) conic fitting problem always has a solution: for any data points $P_1, \ldots, P_n$ one can always find a best fitting object $S_{\text{best}} \in \bar{\mathbb{M}}_{\text{Q}}$ that minimizes the sum of squares of the distances to $P_1, \ldots, P_n$. One has to keep in mind, though, that the best fitting object may be a line, or a pair of parallel lines, or a pair of intersecting lines, rather than a conic.

THEOREM 2.4. *Let B be a given compact set containing all the data points. The 'enlarged' set Ω of ellipses, parabolas, hyperbolas, lines (including rays and line segments, singletons and opposite rays), pairs of parallel lines, pairs of intersecting lines crossing B is compact.*

See Appendix for the proof of the theorem.

Remark

In the last two equations, (2.35) and (2.36), we could have ignored lines based on the redundancy principle (see section 2.4) because every line is a subset of a pair of parallel or intersecting lines. However, we choose not to do that. A single line is a much simpler object than a pair of lines, so it is convenient to have a single line as the best fitting object, whenever possible.

2.9. Sufficiency and deficiency model

The secondary objects (lines, parallel lines, etc) are added to the model collection merely for the purpose of ensuring existence of the best fit. Suppose one fits circles to observed points. In order to close up the model collection of circles, one has to add lines to it. So circles are primary objects, and lines - secondary objects. The best fitting object to a given set of data points then will be a circle, or occasionally a line (which is an less desirable result in a practical applications). Fortunately, the probability of the best fit being a line has zero probability if the data points have continuous distribution (see [**12**] (see Theorem 8 on page 68 there)). However, for some more complicated model collections (ex, ellipses), there is an nonzero probability of getting secondary object as the best fit.

DEFINITION 2.2. *A collection* $\mathbb{M}$ *of model objects is said to be* sufficient *(for fitting purposes) if the best fitting object exists with probability one, assuming that the data points are independent normally distributed random variables (or more generally, that the coordinates of the data points have a joint probability density function). Otherwise, it is* deficient.

For any set of distinct 5 points in a general linear position (which means that no three points are collinear), there exists a unique quadratic curve (conic) passing through all of them (i.e., interpolates them);[**?**]. That conic may be an ellipse, a parabola, a hyperbola. If 5 points are not in general linear position (i.e., at least three of them are collinear), then they can be interpolated by a degenerate conic (a pair of lines).

Probabilistic approach for $n = 5$ points Let us check the most trivial case of fitting five distinct points (The analysis of samples of $n > 5$ points requiring more sophisticated numerical tests done by Hui Ma)). If all five points are interpolated by an ellipse, then it is obviously the best fit. But if the five points are interpolated by secondary objects (lines, parallel lines and parabola) in the closure of model

collection of all ellipses, we have a unwanted event: a secondary object provides the best fit. This, however, occurs with probability zero, so it is not a real concern. But what if interpolating conic is a hyperbola or a pair of intersecting lines? We found that there is no local minimum in the set of ellipses. More precisely, for any ellipse $\mathbb{E}$, there always exists a $\mathbb{E}'$ such that $\mathcal{F}(\mathbb{E}') < \mathcal{F}(\mathbb{E})$ (see section A.1 in appendix). It has been proven that the best fit should exist in the closure of set of all ellipses. So the best fit has to be a secondary object. The numerical experiment shows that random points were interpolated by an ellipse with probability 22% and by a hyperbola with probability 78% (By Hui Ma) Thus the best fitting object S_{best} is now a secondary one (a parabola, or a line, or a pair of parallel liner), i.e., an unwanted event occurs. And this really happens with a positive probability.

Conclusion The collection of ellipses is not sufficient for fitting purposes. This means that there is a real chance that for a given set of data points no ellipse could be selected as the best fit to the points, i.e., the ellipse fitting problem would have no solution. More precisely, for any ellipse E there will be another ellipse E' providing a better fit, in the sense $\mathcal{F}(E') < \mathcal{F}(E)$. If one constructs a sequence of ellipses that fit the given points progressively better and on which the objective function $\mathcal{F}$ converges to its infimum, then those ellipses will grow in size and converge to something different than an ellipse. Most likely, they will converge to a parabola.

Speaking informally, whenever the best fitting ellipse fails to exist, the ellipse fitting procedure attempts to move beyond the collection of ellipses, and ends up on the border of that collection... Then it returns a secondary object (a parabola or a pair of lines). In a sense, the scope of the collection of ellipses is too narrow for fitting purposes. One may say that this collection is seriously deficient, or badly incomplete for fitting purposes. It calls for a substantial extension.

2.10. Megaspace on specific model collections

In Section 2.7 we defined a multidimensional set ("megaset") $\mathfrak{M}(\mathbb{M})$, which corresponds to a given collection $\mathbb{M}$ of model objects. We also showed that the task of finding the best fitting object $S_{\text{best}} \in \mathbb{M}$ to a given set of points $P_1, \ldots, P_n$ is equivalent to projecting the "megapoint" $\mathcal{P}$ corresponding to $P_1, \ldots, P_n$ onto the "megaset" $\mathfrak{M}(\mathbb{M})$. Here we describe the "megasets" for several model collections.

Megaset for Lines Let $\mathbb{M}_{\text{L}}$ consist of all lines in $\mathbb{R}^2$. The corresponding "megaset" $\mathfrak{M}(\mathbb{M}_{\text{L}}) \subset \mathbb{R}^{2n}$ is described by Malinvaud (see Chapter 10 in [**24**]) and Chernov (see Section 1.5 and Section 3.4 in [**12**]). A point (we also call it "megapoint") $(x_1, y_1, \ldots, x_n, y_n)$ belongs to $\mathfrak{M}(\mathbb{M}_{\text{L}})$ if and only if all the n planar points (x_1, y_1), …, (x_n, y_n) belong to one line (i.e., they are collinear). This condition can be expressed by $C_{n,3}$ algebraic relations:

$$\det \begin{bmatrix} x_i - x_j & y_i - y_j \\ x_i - x_k & y_i - y_k \end{bmatrix} = 0 \tag{2.37}$$

for all $1 \le i < j < k \le n$. Each of these relations means that the three points (x_i, y_i), (x_j, y_j), and (x_k, y_k) are collinear. All of these relations together mean that all the n points (x_1, y_1), …, (x_n, y_n) are collinear.

Note that $\mathbb{M}_{\text{L}}$ is specified by $n - 2$ independent relations, hence it is an $(n + 2)$-dimensional manifold (algebraic variety) in $\mathbb{R}^{2n}$. The relations (2.37) are quadratic, so $\mathbb{M}_{\text{L}}$ is a quadratic surface. It is closed in topological sense, hence the problem of finding the best fitting line always has a solution.

Megaset for Circles

Let $\mathbb{M}_{\text{C}}$ consist of all circles in $\mathbb{R}^2$. The corresponding "megaset" $\mathfrak{M}(\mathbb{M}_{\text{C}}) \subset \mathbb{R}^{2n}$ is described by Chernov (see Section 1.5 and Section 3.4 in [**12**]). A "megapoint" $(x_1, y_1, \ldots, x_n, y_n)$ belongs to $\mathfrak{M}(\mathbb{M}_{\text{C}})$ if and only if all the n planar points (x_1, y_1), …, (x_n, y_n) belong to one circle (in such a case we will say that these points are *cocircular*). In that case all these points satisfy one quadratic equation of a special

type:

$$A(x^2+y^2)+Bx+Cy+D=0. \tag{2.38}$$

This condition can be expressed by $C_{n,4}$ algebraic relations:

$$\det\begin{bmatrix} x_i-x_j & y_i-y_j & x_i^2-x_j^2+y_i^2-y_j^2 \\ x_i-x_k & y_i-y_k & x_i^2-x_k^2+y_i^2-y_k^2 \\ x_i-x_m & y_i-y_m & x_i^2-x_m^2+y_i^2-y_m^2 \end{bmatrix}=0 \tag{2.39}$$

for $1 \le i < j < k < m \le n$. Each of these relations means that the four points (x_i, y_i), (x_j, y_j), (x_k, y_k), and (x_m, y_m) satisfy one quadratic equation of type (2.38), i.e., they are either cocircular or collinear. All of these relations together mean that all the n points (x_1, y_1), ..., (x_n, y_n) satisfy one quadratic equation of type (2.38), i.e., they all are either cocircular or collinear. Therefore the relations (2.39) describe the union $\mathfrak{M}(\mathbb{M}_{\mathrm{C}}) \cup \mathfrak{M}(\mathbb{M}_{\mathrm{L}})$.

Relation between the megaset for Circles and the megaset for Lines

The determinant in (2.39) is a polynomial of degree four, and $\mathfrak{M}(\mathbb{M}_{\mathrm{C}}) \cup \mathfrak{M}(\mathbb{M}_{\mathrm{L}})$ is an $(n+3)$-dimensional algebraic variety (manifold) in $\mathbb{R}^{2n}$ defined by quadratic polynomial equations. Note that the dimension of $\mathfrak{M}(\mathbb{M}_{\mathrm{C}}) \cup \mathfrak{M}(\mathbb{M}_{\mathrm{L}})$ is one higher than that of $\mathfrak{M}(\mathbb{M}_{\mathrm{L}})$, i.e.

$$\dim(\mathfrak{M}(\mathbb{M}_{\mathrm{C}}) \cup \mathfrak{M}(\mathbb{M}_{\mathrm{L}})) = \dim\mathfrak{M}(\mathbb{M}_{\mathrm{L}}) + 1. \tag{2.40}$$

A closer examination shows that $\mathfrak{M}(\mathbb{M}_{\mathrm{L}})$ plays the role of the boundary of $\mathfrak{M}(\mathbb{M}_{\mathrm{C}})$, i.e., $\mathfrak{M}(\mathbb{M}_{\mathrm{C}})$ terminates on $\mathfrak{M}(\mathbb{M}_{\mathrm{L}})$. The megaset $\mathfrak{M}(\mathbb{M}_{\mathrm{C}})$ is not closed, but if we add its boundary $\mathfrak{M}(\mathbb{M}_{\mathrm{L}})$ to it, it will become closed.

Megasets for Ellipses and other quadratic curves

Let $\mathbb{M}_{\mathrm{E}}$ consist of all ellipses in $\mathbb{R}^2$. The corresponding "megaset" $\mathfrak{M}(\mathbb{M}_{\mathrm{E}}) \subset \mathbb{R}^{2n}$ can be described in a similar manner as above. A point $(x_1, y_1, \ldots, x_n, y_n)$ belongs in $\mathfrak{M}(\mathbb{M}_{\mathrm{E}})$ if and only if all the n planar points (x_1, y_1), ..., (x_n, y_n) belong to one ellipse (in such a case we will say that these points are *coelliptical*). In that case all

these points satisfy one quadratic equation of a general type:

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0. \tag{2.41}$$

This equation actually means that the points belong to one conic (either regular or degenerate). This condition can be expressed by $C_{n,6}$ algebraic relations:

$$\det \begin{bmatrix} x_i - x_j & y_i - y_j & x_i^2 - x_j^2 & y_i^2 - y_j^2 & x_i y_i - x_j y_j \\ x_i - x_k & y_i - y_k & x_i^2 - x_k^2 & y_i^2 - y_k^2 & x_i y_i - x_k y_k \\ x_i - x_m & y_i - y_m & x_i^2 - x_m^2 & y_i^2 - y_m^2 & x_i y_i - x_m y_m \\ x_i - x_l & y_i - y_l & x_i^2 - x_l^2 & y_i^2 - y_l^2 & x_i y_i - x_l y_l \\ x_i - x_r & y_i - y_r & x_i^2 - x_r^2 & y_i^2 - y_r^2 & x_i y_i - x_r y_r \end{bmatrix} = 0 \tag{2.42}$$

for $1 \le i < j < k < m < l < r \le n$. Each of these relations means that the six points (x_i, y_i), (x_j, y_j), (x_k, y_k), (x_m, y_m), (x_l, y_l), and (x_r, y_r) satisfy one quadratic equation of type (2.41), i.e., they belong to one conic (either regular or degenerate). All of these relations together mean that all the n points (x_1, y_1), ..., (x_n, y_n) satisfy one quadratic equation of type (2.41), i.e., they all belong to one conic (regular or degenerate). Therefore the relations (2.42) describe a much larger megaset $\mathfrak{M}(\mathbb{M}_{\mathrm{Q}})$ corresponding to the collection of all quadratic curves, regular and degenerate, i.e.,

$$\mathbb{M}_{\mathrm{Conics}} = \mathbb{M}_{\mathrm{E}} \cup \mathbb{M}_{\mathrm{H}} \cup \mathbb{M}_{\cup} \cup \mathbb{M}_{\mathrm{L}} \cup \mathbb{M}_{\parallel} \cup \mathbb{M}_{\times} \tag{2.43}$$

where $\mathbb{M}_{\mathrm{H}}$ denotes the collection of all hyperbolas, and other notation was introduced in section 2.8, in which we showed that $\mathbb{M}_{\mathrm{Q}}$ was topologically closed.

Relation between the megasets corresponding to different types of quadratic curve

The determinant in (2.42) is a polynomial of the eighth degree, and $\mathfrak{M}(\mathbb{M}_{\mathrm{Q}})$ is a closed $(n+5)$-dimensional algebraic manifold in $\mathbb{R}^{2n}$. It is mostly made of two big megasets: $\mathfrak{M}(\mathbb{M}_{\mathrm{E}})$ and $\mathfrak{M}(\mathbb{M}_{\mathrm{H}})$, they both are $(n+5)$-dimensional. Other megasets listed in the decomposition (2.43) have smaller dimensions and play the role of the boundaries of the bigger megasets $\mathfrak{M}(\mathbb{M}_{\mathrm{E}})$ and $\mathfrak{M}(\mathbb{M}_{\mathrm{H}})$.

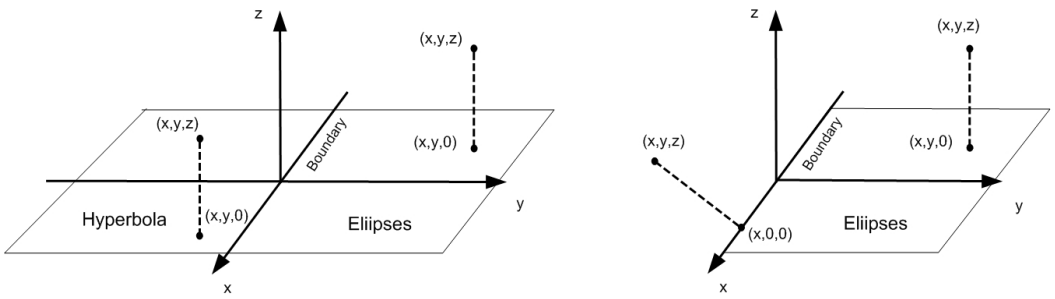


Figure 2.10: Projecting a set of points onto the megaspace of ellipses or general quadratic curves

The structure of the megaset $\mathfrak{M}(\mathbb{M}_{\text{Conics}})$ is schematically illustrated here, where it is shown as the xy plane $\{z = 0\}$ in the 3D space (which plays the role of the megaspace $\mathbb{R}^{2n}$). The positive half-plane $H_+ = \{y > 0, z = 0\}$ represents the elliptic megaset $\mathfrak{M}(\mathbb{M}_{\text{E}})$, and the negative half-plane $H_- = \{y < 0, z = 0\}$ represents the hyperbolic megaset $\mathfrak{M}(\mathbb{M}_{\text{H}})$. The x-axis $\{y = z = 0\}$ separating these two half-planes represents all the lower-dimensional megasets $\mathfrak{M}(\mathbb{M}_{\cup} \cup \mathbb{M}_{\text{L}} \cup \mathbb{M}_{\parallel} \cup \mathbb{M}_{\times})$ in the decomposition (2.43). The real structure of $\mathfrak{M}(\mathbb{M}_{\text{Conics}})$ is much more complicated, but our simplified picture still shows its most basic features.

Sufficiency and deficiency illustrated

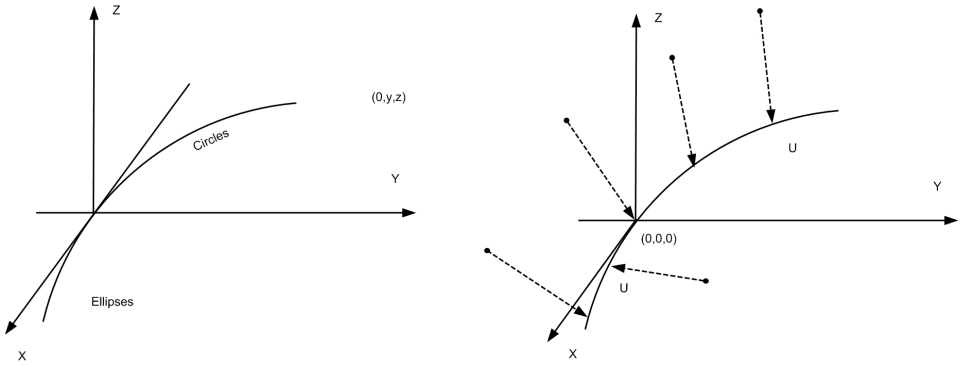
Now recall that finding the best fitting solution corresponds to an orthogonal projection of the given megapoint $\mathcal{P}$ in the megaspace $\mathbb{R}^{2n}$ (in our illustration, it would be a point $(x, y, z) \in \mathbb{R}^3$) onto the megaset $\mathfrak{M}(\mathbb{M}_{\text{Conics}})$ (in our illustration - onto the xy plane). Then the point (x, y, z) is simply projected onto $(x, y, 0)$. What are the chances that the footpoint of the projection corresponds to the "boundary" objects $\mathbb{M}_{\cup} \cup \mathbb{M}_{\text{L}} \cup \mathbb{M}_{\parallel} \cup \mathbb{M}_{\times}$ (i.e., to the secondary objects, in terms of Section 2.9) Clearly, only the points of the xz plane $\{y = 0\}$ are projected onto the line $\{y = z = 0\}$. If the point $(x, y, z) \in \mathbb{R}^3$ is selected randomly with an absolutely continuous distribution (which has a probability density), then a point on the xz plane would be chosen with probability zero. This fact illustrates the sufficiency of the model collection of

non-degenerate conics (even the sufficiency of ellipses and hyperbolas alone, without parabolas): the best fitting object will be a secondary object with probability zero.

But what if our model collection consists of ellipses only, without hyperbolas? Then in our illustration, the corresponding megaset would be the positive half-plane $H_+ = \{y > 0, z = 0\}$. Finding the best fitting ellipse would correspond to an orthogonal projection of the given point $(x, y, z) \in \mathbb{R}^3$ onto the positive half-plane $H_+ = \{y > 0, z = 0\}$. Now if the given point (x, y, z) has a positive y-coordinate, then it is projected onto $(x, y, 0)$, as before, and we get the desired best fitting ellipse. But if it has a negative y-coordinate, then it is projected onto $(x, 0, 0)$, which is on the *boundary* of the half-plane, so we get a boundary footpoint, i.e., a secondary object will be the best fit. We see that all the points (x, y, z) with $y < 0$ (making a whole half-space!) are projected onto the boundary line, hence for all those points the best fitting ellipse would not exist! This fact clearly illustrates the deficiency of the model collection of ellipses.

Sufficiency of circles versus deficiency of ellipses: a controversy or not? One may wonder: How is it possible that the collection of circles is sufficient (as we proved in section 2.8), while the larger collection of ellipses is not? Indeed every circle is an ellipse, hence the collection of ellipses contains all the circles. So why the sufficiency of circles does not guarantee the sufficiency of the bigger, inclusive collection of ellipses? Well, this seemingly counterintuitive fact can be illustrated, too.

Suppose the megaset $\mathfrak{M}(\mathbb{M}_{\mathrm{C}})$ for the collection of circles is represented by the set $U = \{y = x^2, x \neq 0, z = 0\}$ in our illustration. Note that U consists of two curves (branches of a parabola on the xy plane), both lie in the half-plane $H_+ = \{y > 0, z = 0\}$ that corresponds to the collection of ellipses. So the required inclusion $U \subset H_+$ does take place. The two curves making U terminate at the point $(0, 0, 0)$, which does not belong to U, so it plays the role of the boundary of U. Now suppose a randomly selected point $(x, y, z) \in \mathbb{R}^3$ is to be projected onto the set U. What are the chances



that its projection will end up on the boundary of U, i.e., at the origin $(0, 0, 0)$? It is not hard to see (and prove by elementary geometry) that only points on the yz plane may be projected onto the origin $(0, 0, 0)$ (and not even all of them; points with large positive y-coordinates would be projected onto some interior points of U). So the chance that the footpoint of the projection ends up at the boundary of U is zero. This illustrates the sufficiency of the smaller model collection of circles, despite the deficiency of the larger model collection of ellipses (which we have seen above).

Our analysis of the existence of the best fitting object is now complete. We proceed to the uniqueness issue in section 3.1.

CHAPTER 3

UNIQUENESS OF THE BEST FIT

In the previous chapter we have resolved the issue of existence of the best fitting object, in a general setting and for specific classes of models. Here we start addressing the uniqueness issue. Is the best fitting object unique? That is, does the objective function take a unique global minimum? While for typical data sets which are randomly generated, the best fit is unique, there are exceptions. We will deal with three different popular models: line, circle and ellipse. We begin with the simplest fitting model.

3.1. Uniqueness of the best fitting line

In the previous sections we have resolved the issue of existence of the best fitting object, in a general setting and for specific classes of models. Here we start addressing the uniqueness issue. Is the best fitting object unique? That is, does the objective function take a unique global minimum?

Lines

We begin with lines, which are the simplest model objects on our agenda. The uniqueness of the best fitting line has been studied long ago, and a recent summary can be found in Sections 2.2 and 2.3 of [**12**]. We present that summary below.

Sample means and centroid

Given data points (x_1, y_1), ..., (x_n, y_n) we denote by $\bar{x}$ and $\bar{y}$ the sample means

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{and} \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i. \tag{3.1}$$

The point $(\bar{x}, \bar{y})$ is called the center of mass, or the "centroid" of the given data set.

Scatter matrix

We also denote by

$$s_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2$$
$$s_{yy} = \sum_{i=1}^{n} (y_i - \bar{y})^2$$
$$s_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

the components of the so called "scatter matrix"

$$\mathbf{S} = \begin{bmatrix} s_{xx} & s_{xy} \\ s_{xy} & s_{yy} \end{bmatrix}, \tag{3.2}$$

which characterizes the spread of the data set about its centroid $(\bar{x}, \bar{y})$.

This matrix is symmetric, so its eigenvectors are perpendicular to each other. It is also positive-semidefinte, so its eigenvalues are non-negative numbers.

Scattering ellipse

The scatter matrix $\mathbf{S}$ is related to the "scattering ellipse", which is defined by equation

$$\begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix}^T \mathbf{S}^{-1} \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} = n - 1.$$

Its center is the centroid $(\bar{x}, \bar{y})$. Its axes are spanned by the eigenvectors of the scatter matrix $\mathbf{S}$. The major axis is spanned by the eigenvector corresponding to the larger eigenvalue. The minor axis is spanned by the eigenvector corresponding to the smaller eigenvalue. The lengths of its axes are the square roots of the eigenvalues of $\mathbf{S}$.

Next we find the best fitting line following chapter 2 in [**12**]. We will describe lines in the xy plane by equation

$$Ax + By + C = 0 \tag{3.3}$$

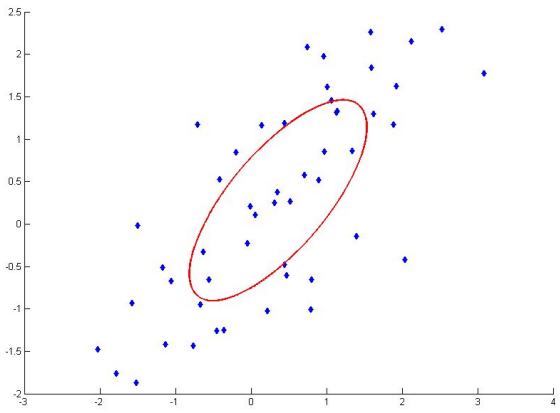


Figure 3.1: Randomly generated data points and the scattering ellipse

where A, B, C are the parameters of the line. The distance from a point $P = (x, y)$ to a line L given by (3.3) is

$$\text{dist}(P_i, L) = \frac{|Ax_i + By_i + C|}{A^2 + B^2} \tag{3.4}$$

Now the best fitting line can be found by minimizing the objective function

$$\mathcal{F}(A, B, C) = \frac{1}{A^2 + B^2} \sum_{1}^{n} (Ax_i + By_i + C)^2 \tag{3.5}$$

The parameters (A, B, C) need only be specified up to a scalar multiple. Thus we can impose a constraint, for example

$$A^2 + B^2 = 1 \tag{3.6}$$

With this constraint, the formula for the objective function simplifies to

$$\mathcal{F}(A, B, C) = \sum_{1}^{n} (Ax_i + By_i + C)^2 \tag{3.7}$$

Since the parameter C is unconstrained, we can eliminate it by minimizing (3.7) with respect to C while holding A and B fixed. Solving the equation $\partial\mathcal{F}/\partial C = 0$ gives us

$$C = -A\overline{x} - B\overline{y} \tag{3.8}$$

In particular, we see that the best fitting line always passes through the centroid $(\overline{x}, \overline{y})$ of the data set. Eliminating C from (3.7) gives

$$\mathcal{F}(A,B) = \sum_{1}^{n} [A(x_i - \overline{x})^2 + B(y_i - \overline{y})^2]^2 \tag{3.9}$$
$$= S_{xx}A^2 + 2S_{xy}AB + S_{yy}B^2$$

or in matrix form

$$\mathcal{F}(H) = H^T S H \tag{3.10}$$

where $H = (A, B)$ denotes the parameter vector. Minimizing (3.10) subject to the constraint $||H|| = 1$ is a simple problem of the matrix algebra: its solution is the eigenvector of the scatter matrix S corresponding to the smaller eigenvalue. Observe that the parameter vector H is orthogonal to the line (3.3), thus the line itself is parallel to the other eigenvector. In addition, it passes through the centroid, hence it is the major axis of the scattering ellipse. The above observations are summarized as follows:

Main facts

THEOREM 3.1. *The best fitting line $Ax + By + C = 0$ always passes through the centroid, i.e., $A\bar{x} + B\bar{y} + C = 0$. It coincides with the major axis of the scattering ellipse.*

For typical data sets, the above procedure leads to a unique best fitting line. But there are certain exceptions. If the two eigenvalues of S coincide, then every vector $H \neq 0$ is its eigenvector and the function $\mathcal{F}(A, B)$ is actually constant on the unit circle $\| H \| = 1$ In that case all the lines passing through the centroid of the data minimize $\mathcal{F}$; hence the problem has multiple (infinitely many) solutions. This happens if and only if S is a scalar matrix, i.e.

$$s_{xx} = s_{yy} \quad \text{and} \quad s_{xy} = 0 \tag{3.11}$$

We emphasize that the orthogonal regression line is not unique if and only if both equations in (3.11) hold. The above observations are summarized as follows:

THEOREM 3.2. *The best fitting line is not unique if and only if the eigenvalues of the scatter matrix* $\mathbf{S}$ *coincide, so that the scattering ellipse becomes a circle. In that case* every *line passing through the centroid* $(\bar{x}, \bar{y})$ *is a best fitting line.*

Dichotomy Thus we have a dichotomy:

- either there is a 'single' best fitting line,
- or there are 'infinitely many' best fitting lines.

In the latter case, the whole bundle of lines passing through the centroid $(\bar{x}, \bar{y})$ are best fitting lines.

Examples A simple example of a data set for which there are multiple best fitting lines is n points placed at the vertices of a regular polygon with n vertices (n-gon). Rotating the data set around its center by the angle $2\pi/n$ takes the data set back to itself. So if there is one best fitting line, then by rotating it through the angle $2\pi/n$ we get another line that fits equally well. Thus the best fitting line is not unique.

It is less obvious (but true, according to Theorem 3.2 above) that every line passing through the center of our regular polygon is a best fitting line; they all minimize the objective function.

Data points placed at vertices of a regular polygon seem like a very exceptional situation. However multiple best fitting lines are much more common. The following is true:

THEOREM 3.3. *Given any data points* (x_1, y_1), ..., (x_n, y_n) *we can always move one of them so that the new data set will admit multiple best fitting lines. Precisely, there are always* x_n' *and* y_n' *such that the set* (x_1, y_1), ..., (x_{n-1}, y_{n-1}), (x_n', y_n') *admits multiple best fitting lines.*

In other words, the $n-1$ points can be placed arbitrarily, without any regular pattern whatsoever, and then we can add just one extra point so that the set of all n points will admit multiple best fitting lines, i.e., will satisfy (3.11).

Still, the existence of multiple best fitting lines is a very unlikely event in probabilistic terms. If data points are sampled randomly from an absolutely continuous probability distribution, then this event occurs with probability zero. Indeed, equations (3.11) specify a subsurface (submanifold) in the $2n$ dimensional space with coordinates $x_1, y_1, \ldots, x_n, y_n$, That submanifold has zero volume, hence for any absolutely continuous probability distribution its probability is zero.

However, if the data points are obtained from a digital image (say, they are pixels on a computer screen), then the chance of having (3.11) may no longer be negligible and may have to be reckoned with. For instance, a simple configuration of 4 pixels making a 2×2 square satisfies (3.11), and thus the orthogonal fitting line is not uniquely defined.

Next we turn to circles in Section **Uniqueness of the best fitting circle**.

3.2. Uniqueness of the best fitting circle

Introduction

After we have seen in Section (3.1) that the simplest fitting problem - involving lines - had multiple solutions, it may not be too surprising to find out that more complicated problems also have multiple solutions. Here we demonstrate this for circles.

Unfortunately, we cannot describe all data sets for which the best fitting circle is not unique in the same comprehensive manner as we did it for lines in the last Section. We can only give some examples of such data sets.

Main idea: rotational symmetry All the known examples are based on the "rotational symmetry" of the data set. We already used this idea in the last Section. Suppose the data set can be rotated around some point O through the angle $2\pi/k$

for some $k \geq 2$, and after the rotation it comes back to itself. Then, if there is a best fitting circle, rotating it around O through the angle $2\pi/k$ would give us another circle that would fit the data set equally well. This is how we get more than one best fitting circle.

This is a nice idea but it breaks down instantly if the center of the best fitting circle happens to coincide with the center of rotation O. Then we would rotate the circle around its own center, hence we would get the same circle again. Thus one has to construct a rotationally symmetric data set more carefully to avoid best fitting circles centered on the natural center of symmetry of the set.

Example by Nievergelt

The earliest and simplest example was given by Nievergelt on pages 260-261 in [**27**]. He chose $n = 4$ data points as follows:

$$(0, 0), \quad (0, 2), \quad (\sqrt{3}, -1), \quad (-\sqrt{3}, -1)$$

Three last points are at the vertices of an equilateral triangle centered on $(0, 0)$. So the whole set can be rotated around the origin $(0, 0)$ through the angle $2\pi/3$ and it will come back to itself.

Nivergelt claimed that the best fitting circle has center $(0, -3/4)$ and radius $R = 7/4$. This circle passes through the last two data points and cuts right in the middle between the first two. So the first two points are at distance $d = 1$ from that circle and the last two are right on it (their distance from the circle is zero). Thus the objective function is

$$\mathcal{F}_1 = 1^2 + 1^2 + 0^2 + 0^2 = 2. \tag{3.12}$$

It is easy to believe that Nivergelt's circle is the best, indeed, as any attempt to perturb its center or radius would only make the fit worse (the objective function would grow). However a complete mathematical proof of this claim would be perhaps prohibitively difficult.

Proof of multiplicity in Nievergelt's example

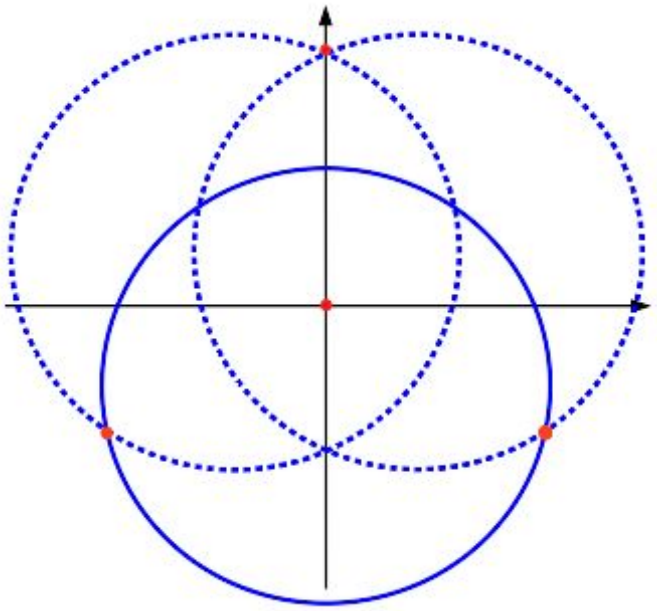


Figure 3.2: Nievergelt's example: Four data points (red) Three fitting circles (blue)

The goal is to show that "there are" multiple best fitting circles (without finding them explicitly). And the multiplicity here can be proven fully in three easy steps.

Step 1. According to our section (2.8), for every data set the best fitting object exists, which may be a circle or a line. If the best object is a circle, then its center is either at $(0, 0)$ or elsewhere. So we have three possible cases: The best fitting object is a line. The best fitting object is a circle centered on $(0, 0)$. The best fitting object is a circle with a center different from $(0, 0)$. In the last case our rotational symmetry will work, as explained above, and prove the multiplicity of the best fitting circle. So we need to rule out the first two cases.

Step 2. Consider any circle of radius R centered on $(0, 0)$. It is easy to see that the respective objective function is

$$\mathcal{F} = R^2 + 3(2 - R)^2 = 4R^2 - 12R + 12.$$

Its minimum is attained at $R = 3/2$ and its minimum value is

$$\mathcal{F}_2 = (3/2)^2 + 3(1/2)^2 = 3. \tag{3.13}$$

This is larger than $\mathcal{F}_1 = 2$ in (3.15). Thus circles centered on the origin cannot compete with Nievergelt's circle and should be ruled out.

Step 3. Consider all lines. As we have seen in Section (3.1), for rotationally symmetric data sets all the best fitting lines pass through the center. All of those lines fit equally well. Taking the x axis, for example, it is easy to see that the corresponding objective function is

$$\mathcal{F}_3 = 2^2 + 1^2 + 1^2 + 0^2 = 6. \tag{3.14}$$

This is greater than $\mathcal{F}_1 = 2$ in (3.15) and even greater than $\mathcal{F}_2 = 3$ in (3.16). Thus lines are even less competitive than circles centered on the origin, so they are ruled out as well. The proof is finished.

Conclusion. The best fitting circle has a center different from $(0, 0)$. Thus by rotating this circle through the angles $2\pi/3$ and $4\pi/3$ we get two more circles that fit the data equally well. So the circle fitting problem has three distinct solutions. The alledged best fitting circles are shown in our illustration.

Other examples

After Nievergelt's example, two other papers presented, independently, similar examples of non-unique circle fits.

Chernov and Lesort [**15**] used a perfect square, instead of Nievergelt's regular triangle. They placed four points at the vertices of the square, and another 4 points at its center, so the data set consisted of $n = 8$ points total. Then they used the above strategy to prove that at least four different circles achieve the best fit.

Zelniker and Clarkson [**31**] used a regular triangle again, placed three points at its vertices and three more points at its center (so that the data set consisted of $n = 6$ points). Then they showed that at least three different circles achieve the best fit.

These examples lead to an interesting fact that may seem rather counterintuitive. Let C be a circle of radius R with center O. Let us place a large number of data points on C and a single data point at the center O. Suppose the points on C are placed uniformly (say at the vertices of a regular polygon). Then it seems like C is

an excellent candidate for the best fitting circle it interpolates all the data points and misses only at O, so $\mathcal{F} = R^2$. It is hard to imagine that any other circle or line can do any better.

However, a striking fact proved by Nievergelt ([**28**], Lemma 7) says that the center of the best fitting circle cannot coincide with any data point. Therefore in our example C cannot be the best fitting circle. Hence some other circle with center $O' \neq O$ fits the data set better. And again, rotating the best circle about O gives other best fitting circles, so those are not unique.

How unusual are multiple circle fits?

Rotationally symmetric data sets described above are clearly exceptional; small perturbations of data points easily destroy the symmetry. But there are probably many other data sets, without any symmetries, that admit multiple circle fits, too. We believe that they are all unusual and can be easily destroyed by small perturbations. Below is our argument.

Suppose a set of data points $P_1, \ldots, P_n$ admits two best circle fits, and denote those circles by C_1 and C_2. First consider a simple case: C_1 and C_2 are concentric, i.e., have a common center, O. Let D_i denote the distance from the point P_i to the center O. By direct inspection, for any circle of radius R centered on O the objective function is

$$\mathcal{F} = \sum_{i=1}^{n} (R - D_i)^2 = nR^2 - 2R \sum_{i=1}^{n} D_i + \sum_{i=1}^{n} D_i^2.$$

This is a quadratic polynomial in R, so it cannot have two distinct minima. So the two best fitting circles cannot be concentric.

Now suppose the circles C_1 and C_2 are not concentric, i.e., they have distinct centers, O_1 and O_2. Let L denote the line passing through O_1 and O_2. Note that the data points cannot be all on the line L (because if the data points were collinear, the best fit would be achieved by the interpolating line and not by two circles). So there exists a point P_i that does not lie on the line L. Hence we can move it slightly toward the circle C_1 but away from the circle C_2. Then the objective function $\mathcal{F}$ changes

slightly, and it will decrease at one minimum (on C_1) and increase at the other (on C_2). This will break the tie and ensure the uniqueness of the global minimum.

We proceed to the section **Uniqueness of the best fitting ellipse**.

3.3. Uniqueness of the best fitting ellipse

Introduction

In Section (3.2) we showed several examples of data sets for which the best fitting circle was not unique. It should not be surprising now that data sets exist for which the best fitting ellipse is not unique either.

Rotational symmetry

The main idea of all known examples of multiple fits is the rotational symmetry of the data set, as described in section (3.2). Suppose the data set can be rotated around some point O through the angle $2\pi/k$ for some $k \geq 2$ and after the rotation it comes back to itself. Then if there is a best fitting object, rotating it around O through the angle $2\pi/k$ would give us another object that would fit the data set equally well.

In fact the above example obeys this principle: rotating a perfect square around its center through $\pi/2$ brings it back to itself. Likewise, rotating a perfect square lattice of $N \times N$ points around its center through $9\pi/2$ brings it back to itself. Thus, rotating one ellipse around its center by $\pi/2$ produces another best fitting ellipse.

Nievergelt-type example

In section (3.2) we described perhaps the simplest possible example of a multiple circle fit, published by Nievergelt on pages $260 - 261$ in [**27**]. It consisted of $n = 4$ data points: three were placed at vertices of an equilateral triangle, and the fourth one - at its center.

Recall that a circle has three independent parameters, but ellipse - five. So it is natural to generalize Nievergelt's example by placing five data points at vertices of a regular pentagon, and the sixth one - at its center. Thus we have $n = 6$ data points:

$$(0,0), \quad (0,2), \quad \big(\pm 2\cos(\pi/10), 2\sin(\pi/10)\big), \quad \big(\pm 2\cos(3\pi/10), -2\sin(3\pi/10)\big).$$

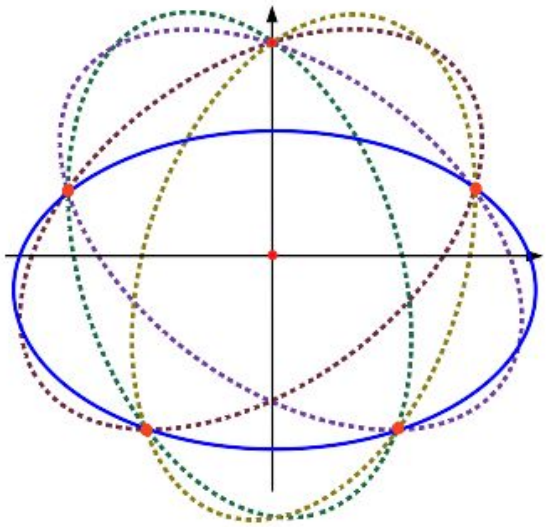


Figure 3.3: Nievergelt-type example: Six data points (red) Five fitting ellipses colors

We strongly believe that the best fitting ellipse passes through the last four data points and the point $(0, 1)$. These five points determine the ellipse uniquely. It is obviously symmetric about the y axis, so its major axis is horizontal. This ellipse cuts right in the middle between the first two data point. So those two points are at distance $d = 1$ from that ellipse and the last four are right on it (the distance is zero). Thus the objective function is

$$\mathcal{F}_1 = 1^2 + 1^2 + 0^2 + 0^2 + 0^2 + 0^2 = 2. \tag{3.15}$$

Below we provide a partial proof of our claim that the above ellipse is the best. We also describe a full computer-assisted proof that involves extensive numerical computations.

Lastly, by rotating this ellipse through the angles $2\pi k/5$ for $k = 1, 2, 3, 4$ we get four more ellipses that fit the data equally well. So the ellipse fitting problem has five distinct solutions; see illustration.

Partial proof for our example

We will compare our ellipse to the best fitting circle centered on the origin and the best fitting lines.

Consider any circle of radius R centered on $(0,0)$. It is easy to see that the respective objective function is

$$\mathcal{F} = R^2 + 5(2-R)^2 = 6R^2 - 20R + 20.$$

Its minimum is attained at $R = 5/3$ and its minimum value is

$$\mathcal{F}_2 = (5/3)^2 + 5(1/3)^2 = 10/3 = 3.333... \tag{3.16}$$

This is larger than $\mathcal{F}_1 = 2$ in (3.15). Thus circles centered on the origin cannot compete with our ellipse.

Consider all lines. As we have seen in Section (3.1), for rotationally symmetric data sets all the best fitting lines pass through the center. All of those lines fit equally well. Taking the x axis, for example, it is easy to see that the corresponding objective function is

$$\mathcal{F}_3 = 2^2 + 2\big(2\sin(\pi/10)\big)^2 + 2\big(2\sin(3\pi/10)\big)^2 = 10. \tag{3.17}$$

This is greater than $\mathcal{F}_1 = 2$ in (3.15) and even greater than $\mathcal{F}_2 = 3.333$ in (3.16). Thus lines are even less competitive than circles centered on the origin.

Also, in the ellipse fitting problem, pairs of parallel lines are legitimate model objects, see section (2.8). We examined the fits achieved by pairs of parallel lines. The best fit we found was by two horizontal lines $y = y_1$ and $y = y_2$, where

$$y_1 = \big(2 + 4\sin(\pi/10)\big)/4 \qquad \text{and} \qquad y_2 = -2\sin(3\pi/10).$$

Note that y_1 is the average y-coordinate of the first four points in our sample. Thus the first line is the best fitting line for the first four points, and the second line passes through the last two points. The objective function for this pair of lines is

$$\mathcal{F}_4 = (2-y_1)^2 + 2\big(2\sin(\pi/10) - y_1\big)^2 + (0-y_1)^2 \approx 2.146. \tag{3.18}$$

This is pretty good, better than the best fitting circle in (3.16). But still it is a little worse than the best fitting ellipse in (3.15).

Thus our ellipse fits better than any circle centered on the origin, any line, or any pair of parallel lines. In order to conclude that it is really the best fitting ellipse, we would have to compare it to all other ellipses and parabolas. This task seems prohibitively difficult if one uses only theoretical arguments as above. Instead, we developed a computer-assisted proof as described below.

Computer-assisted proof

The ellipse can be specified by five independent parameters, which can be selected in many different ways(e.x. two axes, coordinates of the center and angle of tilt). Suppose we select those parameters and denote them by $(\theta_1, \dots, \theta_5)$. They vary in some domain (parameter space), which we denote by Θ, so that $(\theta_1, \dots, \theta_5) \in \Theta \subset \mathbb{R}^5$.

Now suppose $\tilde{\boldsymbol{\theta}} = (\tilde{\theta}_1, \dots, \tilde{\theta}_5) \in \Theta$ is a certain point in the parameter space. It corresponds to a certain ellipse, for which we can compute the value of the objective function, $\mathcal{F}(\tilde{\boldsymbol{\theta}})$. This value is expected to be greater than the value corresponding to our best ellipse (3.15), i.e., $\mathcal{F}(\tilde{\boldsymbol{\theta}}) > 2$. We proceed assuming that it is indeed, greater than 2.

Suppose also that we estimate (from above) the partial derivative of the objective function $\mathcal{F}$ with respect to each parameter θ_i. Such an estimate must be obtained theoretically and it must be guaranteed to be valid within a certain interval of the values of θ_i. That is, we must derive an upper estimate

$$|\partial\mathcal{F}/\partial\theta_i| \le M_i \qquad \text{for all} \qquad \theta_i \in (\tilde{\theta}_i - a_i, \tilde{\theta}_i + a_i) \tag{3.19}$$

for some $a_i > 0$. The details of our estimation of partial derivatives of $\mathcal{F}$ are given in section A.3.

THEOREM 3.4. *Suppose we have estimates* (3.19) *and let* $0 < b_i < a_i$ *be some numbers. Then for all parameter values* $(\theta_1, \dots, \theta_5) \in \Theta$ *such that* $\theta_i \in (\tilde{\theta}_i - b_i, \tilde{\theta}_i + b_i)$

for every $i = 1, \dots, 5$, we have the following lower bound on the objective function:

$$\mathcal{F}(\theta_1, \dots, \theta_5) \geq \mathcal{F}(\tilde{\boldsymbol{\theta}}) - \sum_{i=1}^{5} M_i b_i.$$

This theorem easily follows from the Mean Value Theorem in calculus.

Now we need to choose the b_i's small enough so that the above lower bound is greater than 2, i.e.,

$$\mathcal{F}(\tilde{\boldsymbol{\theta}}) - \sum_{i=1}^{5} M_i b_i > 2.$$

Then we can conclude that the objective function $\mathcal{F}$ does not attain its minimum in the domain described by

$$\tilde{\theta}_i - b_i \leq \theta_i \leq \tilde{\theta}_i + b_i \qquad \text{for every} \qquad i = 1, \dots, 5. \tag{3.20}$$

This gives us a little domain (a block) in the parameter space, which is "safe" – there are no ellipses there which could beat our best ellipse (3.15).

Of course, the numbers $b_1, \dots, b_5$ may be very small, and so the safety block (3.20) may be very tiny. But it is just one small step in our computer-assisted proof. Then we select another point $\tilde{\boldsymbol{\theta}} = (\tilde{\theta}_1, \dots, \tilde{\theta}_5) \in \Theta$ near the current safety block (e.g., on its boundary), construct another safety block, etc.

This process should be continued step by step, until we cover the entire parameter space Θ by small safety blocks. This strategy resembles the work of a minesweeper in a sea cleaning a large area by finding and disabling dangerous mines, one small area at a time.

Of course there are at least five points in the parameter space Θ where the objective function $\mathcal{F}$ does take the value 2 (which we expect to be its global minimum). Thus our procedure slows down in the vicinity of those five points, we will never be able to reach them. So in fact our minesweeping strategy covers the entire parameter space except small vicinities of five points corresponding to the ellipses described above. This is still OK, as the following logical conclusion can be made.

We know for sure that the objective function takes its global minimum in the vicinity of one of those five points in Θ corresponding to the five ellipses described above. In other words, we know for sure that the best fitting ellipse exists and is quite close to one of the five ellipses described above. Then the rotational symmetry guarantees that there are four more best fitting ellipses obtained by rotation through angles $2k\pi/5$ for $k = 1, 2, 3, 4$. This completes our computer-assisted proof.

CHAPTER 4

PARAMETER SPACE FOR QUADRATIC CURVES (CONICS) ON $\mathbb{S}^5$

In the regression model, parameters for the best fitting conic are estimated based on the minimization of the objective function

$$\mathcal{F}(S) = \sum_{i=1}^{n} \big[\text{dist}(P_i, S)\big]^2, \tag{4.1}$$

where S belongs to our model collection. In this chapter the parameter space is restricted the unit sphere $\mathbb{S}^5$ and investigate the theoretical properties of this reduced space. Let us begin with some basic notations.

4.1. General quadratic equations and algebraic parameters

In the regression model of fitting any conic to points in R^2, there are many types of geometric objects such as ellipse, parabola and hyperbola. To ensure the existence of the best fit, one has to include some degenerate conics (see section (2.8)). The equation for conics has the following form:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{4.2}$$

where A, B, C, D, E, F are real numbers ("parameters" of the conic). If $A = \ldots = F = 0$, then the equation represents the entire plane R^2. This unwanted "conic" should be excluded, so we will always assume that at least one parameter is different from zero, or equivalently $A^2 + B^2 + \ldots + F^2 > 0$.

Since A, B, C, D, E, F are the coefficients of a quadratic polynomial (i.e., an algebraic expression), they are often called "algebraic parameters" of the conic.

Many people use a slightly different form:

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0 \tag{4.3}$$

which we will see later has several advantages (see section (4.4)). In our studies we adopt the form (4.3) and we call the algebraic parameters of the conic the coefficients A, B, C, D, E, F that appear in (4.3).

Parameter vector on the unit sphere $\mathbb{S}^5$ Apparently, the equation for the same conic is unique. For any $t \neq 0$

$$tAx^2 + 2tBxy + tCy^2 + 2tDx + 2tEy + tF = 0 \tag{4.4}$$

represents the same conic as (2.41). The new equation (4.4) has parameters tA, tB, tC, tD, tE, tF. its parameter vector is $t\mathbf{P}$.

To avoid linearly dependent parameter vector, we can use a simple reduction to the parameter space:

$$A^2 + B^2 + C^2 + D^2 + E^2 + F^2 = 1. \tag{4.5}$$

In other words, the parameter vector is always assumed to lie on the unit sphere $\mathbb{S}^5$ and any quadratic curve can be represented by a point on $\mathbb{S}^5$. One could further reduce the $\mathbb{S}^5$ to a half-sphere (a hemisphere) to avoid duplicity caused by two opposite parameter vector $-P$ and P, e.g., by requiring $F \geq 0$. But such a further reduction has little advantage but causes unpleasant technical complications in real applications. So we will work with the sphere $\mathbb{S}^5$ and simply keep in mind that any two diametrically opposite points always represent the same quadratic curve.

4.2. CLASSIFICATION OF CONICS

In the last section we introduce a simple reduction for the algebraic parameter space:$A^2 + B^2 + C^2 + D^2 + E^2 + F^2 = 1$. Thus any quadratic curve can be represented by a point on the unit sphere $\mathbb{S}^5$. The equation for conics has the following form:

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0 \tag{4.6}$$

H	Δ	J	$\Delta \cdot I$	K	F	Conic
> 0	$\neq 0$	< 0				Hyperbola
> 0	$\neq 0$	0				Parabola
> 0	$\neq 0$	> 0	< 0			Ellipse
> 0	$\neq 0$	> 0	> 0			Imaginary ellipse
> 0	0	< 0	0			Intersecting lines
> 0	0	> 0	0			Single point
> 0	0	0	0	< 0		Distinct parallel lines
> 0	0	0	0	> 0		Imaginary parallel lines
> 0	0	0	0	0		Coincident lines
0	0	0	0	< 0		Single line
0	0	0	0	0	$\neq 0$	Poles

Table 4.1: Classification of Conics

To determine the type of conic defined by (4.6), let us consider following matrices and determinants:

$$\Delta = \begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix} \qquad J = \begin{vmatrix} A & B \\ B & C \end{vmatrix} \qquad I = A + C \qquad H = A^2 + B^2 + C^2$$

$$K = \begin{vmatrix} A & D \\ D & F \end{vmatrix} + \begin{vmatrix} C & E \\ E & F \end{vmatrix}$$

The types of conics are classified in terms of the above quantities in the following table 4.1 [**9**]:

Main types of quadratic curves and Imaginary objects The quadratic curves and degenerate ones are represented by equation (4.3) if it has a real solution. The "imaginary object" indicates that the equation (4.3) does not have a real solution in R^2 but corresponds to some specific quadratic curve on the complex plane (e.x

Real non-degenerate	Real degenerate	Empty in $\mathbb{R}^2$
Ellipse $x^2 + y^2 - 1 = 0$	Intersecting line xy=0	Imaginary ellipse $x^2 + y^2 + 1 = 0$
Hyperbola $x^2 - y^2 - 1 = 0$	Parallel lines $x^2 - 1 = 0$	Imaginary parallel lines $x^2 + 1 = 0$
Parabola $x^2 + y = 0$	Single line $x = 0$ or $x^2 = 0$	Poles $1 = 0$ and $-1 = 0$
	Single point $x^2 + y^2 = 0$	

Table 4.2: Examples of Conics

$x^2/3 + y^2/4 = -1$). Thus every parameter vector (A, B, C, D, E, F) (known as one type of algebraic parameters) either represents a conic in R^2 or an imaginary conic which can be considered as am empty solution for our model. Below we list the main types of conics (real and imaginary), with examples of equations representing them: Remark: There are two types of quadratic equations representing single lines:

- Equations with a non-zero quadratic part, such as $x^2 = 0$, represent pairs of coincident lines
- Equations with a non-zero quadratic part, such as $x = 0$, represent single lines

Two poles Any point point on the unit sphere represents a conic (real or imaginary), except two points $(0, 0, 0, 0, 0, 1)$ and $(0, 0, 0, 0, 0, -1)$. They corresponds to the equation:

$$1 = 0 \qquad -1 = 0 \tag{4.7}$$

which have no solution, either real or complex. We call them **North Pole** and **South Pole**. They will play a special role in our analysis.

Geometric dimension and algebraic dimension The "geometric dimension" is the number of parameters needed to specify a geometric figure of the given type.

For example, a single point requires two parameters - its x and y coordinates, so the geometric dimension for the "single point" type is two. A line requires two parameters (say, slope and intercept), so its geometric dimension is also two.

Dimension= 5	Dimension= 4	Dimension= 3 or 2
Ellipse (**E**)	Parabola (**P**)	Parallel lines (**PL**)
Hyperbola (**H**)	Single point (**SP**)	Imaginary parallel lines (**IPL**)
Imaginary ellipse (**IE**)	Intersecting line (**IL**)	Coincident lines (**CL**)
		Single line (**SL**)

Table 4.3: Dimensionality of Conics

A pair of parallel lines needs three parameters - one (common) slope and two intercepts. A pair of intersecting lines needs two parameters for each line, a total of four.

A parabola is completely specified by its directrix and focus. The directrix is a line, so it requires two parameters; the focus is a point, so it takes two more. Thus, the total is four.

An ellipse can be specified by five geometric parameters - the coordinates of its center, the lengths of its axes, and the slope of its major axis. The same applies to hyperbolas.

Empty objects need no geometric parameters.

The "algebraic dimension" of each type of conics characterizes the corresponding set of parameter vectors in the unit sphere $\mathbb{S}^5$.

The sphere $\mathbb{S}^5$ itself is five-dimensional; But its parts may have different dimensionality. Intuitively, the number of dimensions is the minimal number of internal coordinates. Isolated points (like Poles) need no internal coordinates, so their dimension is zero. Lines and curves have dimension one. Planes and surfaces have dimension two, etc. Larger, more complex parts of $\mathbb{S}^5$ have higher dimensions.

The algebraic dimension of each type of conics can be counted by the number of "equality constraints" required for that type. The constraints are listed in the first five columns of the above table. "Inequality constraints", like "> 0", "< 0", or

"$\neq 0$", do not affect the algebraic dimension and should not be counted. But each "pure zero" in the first five columns of the table should be counted.

Each equality constraint (shown as a "pure zero" in the table) reduces the algebraic dimension by one, except the constraint $H = 0$, which reduces the algebraic dimension by 3. Indeed, $H = 0$ implies that $A = 0$, $B = 0$, and $C = 0$, i.e., it enforces three different equality constraints!

The rule is: the algebraic dimension of each type of conics is equal to five minus the number of equality constraints.

This rule works for all types of conics except Poles. In that last case, the combination of $H = 0$ and $K = 0$ implies not only that $A = B = C = 0$, but also $D^2 + E^2 = 0$, hence $D = 0$ and $E = 0$. Thus the constraint $K = 0$, if combined with $H = 0$, enforces two different equality constraints! Now we see that $H = 0$ eliminates three dimensions, after which $K = 0$ eliminates two more, and we obtain $5-3-2 = 0$ algebraic dimensions left.

Typically, the geometric dimension agrees with the algebraic dimension. But there is one exception: the "single point" type has geometric dimension 2 and algebraic dimension 4. While geometrically a point requires two coordinates, analytically its equation involves four degrees of freedom, as we show next. We will show that a single point with coordinates (p, q) is defined by equation

$$a^2(x-p)^2 + b^2(y-q)^2 + 2c(x-p)(y-q) = 0$$

where a and b are arbitrary non-zero numbers and $c^2 < a^2b^2$. Indeed, the above equation can be rewritten as

$$\Big[a(x-p) + \frac{c}{a}\,(y-q)\Big]^2 + \Big[b^2 - \frac{c^2}{a^2}\Big](y-q)^2 = 0.$$

Due to our requirement on c we have $b^2 - \frac{c^2}{a^2} > 0$, so both coefficients are positive. Thus each term in the above equation must be equal to zero. From the second term we get $y = q$ and then from the first we get $x = p$, hence the equation defines a single point, (p, q).

In the above equation all the five parameters a, b, c, p, q are free independent variables, except they must be constrained by the requirement that the resulting parameter vector $(A, B, C, D, E, F)^T$ belongs to the unit sphere $\mathbb{S}^5$. This leaves us with four degrees of freedom.

4.3. Topological space on the unit sphere

Since space of all conics can be represented by points on the unit sphere $\mathbb{S}^5$, each type of conic or imaginary conic hold a corresponding domain on $\mathbb{S}^5$. We will provide a detailed analysis about their topological natures.

Domains on the unit sphere Recall that the conics represented by the equation (4.6) can be classified into 8 different types. In addition, there are two poles: $\mathbf{P}_1 = (0, 0, 0, 0, 0, 1)$ and $\mathbf{P}_{-1} = (0, 0, 0, 0, 0, -1)$ on the sphere ("North Pole" and "South Pole") which have dimension zero. Accordingly, the unit sphere was divided into 10 domains according to types of conics (including imaginary ones and two poles). So we have

$$\mathbb{S}^5 = \mathbb{D}_{\mathrm{E}} \cup \mathbb{D}_{\mathrm{H}} \cup \mathbb{D}_{\mathrm{IE}} \cup \mathbb{D}_{\mathrm{P}} \cup \mathbb{D}_{\mathrm{SP}} \cup \mathbb{D}_{\mathrm{IL}} \cup \mathbb{D}_{\mathrm{PL}} \cup \mathbb{D}_{\mathrm{IPL}} \cup \mathbb{D}_{\mathrm{CL}} \cup \mathbb{D}_{\mathrm{SL}} \cup \{\mathbf{P}_1\} \cup \{\mathbf{P}_{-1}\}, \tag{4.8}$$

where the domains are coded by the names of the conic types, as shown in the table (4.3).

Let us define following functions $\mathbb{S}_5 \to \mathbb{R}$ using defined determinants and matrices in section 4.2 :

$$f_\Delta(\mathbf{P}) = \Delta \quad f_J(\mathbf{P}) = J \quad f_{\Delta I}(\mathbf{P}) = \Delta \cdot I \quad f_K(\mathbf{P}) = K \quad f_H(\mathbf{P}) = H$$

where $\mathbf{P} = A, B, C, D, E, F$. Apparently they are continuous everywhere on R^5 and the domain of points for each type of quadratic conics can be represented as follows:

(4.9)

$$\mathbb{D}_{\mathrm{E}} = f_\Delta^{-1}((-\infty, 0) \bigcup (0, +\infty)) \bigcap f_J^{-1}((0, +\infty)) \bigcap f_{\Delta I}^{-1}((-\infty, 0)) \bigcap f_H^{-1}((0, +\infty))$$

$$\mathbb{D}_{\mathrm{P}} = f_\Delta^{-1}((-\infty, 0) \bigcup (0, +\infty)) \bigcap f_J^{-1}(0) \bigcap f_H^{-1}((0, +\infty)) \tag{4.10}$$

$$\mathbb{D}_{\mathrm{H}} = f_{\Delta}^{-1}((-\infty, 0)\bigcup(0, +\infty))\bigcap f_{J}^{-1}((-\infty, 0))\bigcap f_{H}^{-1}((0, +\infty)) \tag{4.11}$$

Ellipse and Hyperbola

By definition of continuous function in a topological sense, $f_{\Delta}^{-1}((-\infty, 0)\bigcup(0, +\infty))$, $f_{J}^{-1}((0, +\infty))$, $f_{\Delta I}^{-1}((-\infty, 0))$ and $f_{H}^{-1}((-\infty, 0)\bigcup(0, +\infty))$ are open sets whose intersections constitute $\mathbb{D}_{\mathrm{E}}$. Similar reason prove that the openness of $\mathbb{D}_{\mathrm{H}}$. Therefore both $\mathbb{D}_{\mathrm{E}}$ and $\mathbb{D}_{\mathrm{H}}$ are open sets on $\mathbb{S}^5$. This also implies the there is nonzero probability of a random parameter points falling into $\mathbb{D}_{\mathrm{E}}$ and $\mathbb{D}_{\mathrm{H}}$. We will provide more detail in the section (4.4).

Parabola

In expression (4.10), we see that $\mathbb{D}_{\mathrm{P}}$ is represented by the intersection of a closed set $f_{J}^{-1}(0)$ and two open sets. But it is not clear whether $\mathbb{D}_{\mathrm{P}}$ is open or closed. In fact, it is neither open nor closed. First, there exists a sequence of points in $\mathbb{D}_{\mathrm{P}}$ converge to a limit point in $\mathbb{D}_{\mathrm{SL}}$. Consider a sequence below:

$$\mathbf{P}_n = \frac{1}{\sqrt{2n^2+1}}(1, 0, 0, 0, n, n) \qquad n = 1, 2, 3, ... \tag{4.12}$$

As $n \to \infty$, $\mathbf{P}_n$ approaches to a limit point $(0, 0, 0, 0, \sqrt{2}/2, \sqrt{2}/2)$ which corresponds to a horizontal line $\sqrt{2}/2y + \sqrt{2}/2 = 0$. Then $\mathbb{D}_{\mathrm{P}}$ is not a closed set. Next, for any open neighborhood of point $\mathbf{P} \in \mathbb{D}_{\mathrm{P}}$ that satisfies $H > 0$, $\Delta \neq 0$ and $J = 0$, there always exits a point $\mathbf{P}^{'}$ so that $H > 0$, $\Delta \neq 0$ but $J \neq 0$ (either ellipse or hyperbola). Clearly, $\mathbf{P}^{'}$ doesn't belong to $\mathbb{D}_{\mathrm{P}}$, which does not meet the requirement for open sets.

Degenerate conics

In a problem of fitting any quadratic curve , the closed space of models that ensures existence of the minimum of sum of squares of distances consists of hyperbola, ellipse, parabola and all their limiting objects. A quadratic curve may converge to an object of many types. Besides the three major types, the limiting object might be a point, a pair of parallel lines, a ray (one ray or two opposite rays), a line segment, a single line and a pair of intersecting lines. We treat ray and line segment as a part of a full line and include them into the domain of lines $\mathbb{D}_{\mathrm{L}}$ on the unit sphere. The domains of single

points, distinct parallel lines, intersecting lines and coincident lines are characterized by $H > 0$ and $\Delta = 0$ (see section (4.2)). They are neither open nor closed. Indeed, for example, the sequence of unit parameter vectors

$$\mathbf{P}_n = \frac{1}{\sqrt{4n^4 + 2n^2 + 2}} \cdot (1, 0, 1, -n, -n, 2n^2) \in \mathbb{D}_{\mathrm{SP}} \qquad n = 1, 2, 3, \ldots \tag{4.13}$$

corresponding to single points converges to a limit $(0, 0, 0, 0, 0, 1)$ for which $H = 0$. Remember that this is a north pole on the unit sphere, which represents empty set. Thus $\mathbb{D}_{\mathrm{SP}}$ is not closed.

Next, any neighborhood of a point $(\sqrt{2}/2, 0, \sqrt{2}/2, 0, 0, 0) \in \mathbb{D}_{\mathrm{SP}}$ contains a parameter vector $(1-\varepsilon^2/2, 0, 1-\varepsilon^2/2, 0, 0, \varepsilon) \in \mathbb{D}_{\mathrm{IE}}$ ($\varepsilon^2 > 0$ is arbitrarily small). So $\mathbb{D}_{\mathrm{SP}}$ is neither open nor closed. One could also see that the other domains (parallel lines, intersecting lines, coincident lines) are neither open nor closed as well by considering similar examples.

The domains of two coincident lines $\mathbb{D}_{\mathrm{CL}}$ and single lines $\mathbb{D}_{\mathrm{SP}}$ can be combined as a domains $\mathbb{D}_{\mathrm{L}}$ as these two types of objects are geometrically equivalent. For example, $(\sqrt{6}/6, -2\sqrt{6}/6, \sqrt{6}/6, 0, 0, 0)$ and $(0, 0, 0, \sqrt{2}/2, \sqrt{2}/2, 0)$ both correspond to the "same line" $x = y$.

This simple observation proves that $\mathbb{D}_{\mathrm{L}}$ is not closed. Let us consider the sequence of parameter vectors

$$\mathbf{P}_n = (0, 0, 0, \frac{1}{n}, \frac{1}{n}, \sqrt{1 - \frac{2}{n^2}})(\in \mathbb{D}_{\mathrm{L}}) \to (0, 0, 0, 0, 0, 1) \tag{4.14}$$

where $(0, 0, 0, 0, 0, 1)$ represents the north pole. Next, for any neighborhood of $(\sqrt{3}/3, -2\sqrt{3}/3, \sqrt{3}/3, 0, 0, 0)$, there exists a point $(\sqrt{3-\varepsilon}/3, -2\sqrt{3-\varepsilon}/3, \sqrt{3-\varepsilon}/3, 0, 0, \sqrt{6}\varepsilon/3)$ corresponding to two intersecting lines. So $\mathbb{D}_{\mathrm{L}}$ are neither open nor closed.

Imaginary objects and Poles

There are three special domains where the parameter vectors do not describe any real conic because their corresponding quadratic equations have no real solution. They will play special roles in our analysis. Let $\mathbb{D}_{\mathrm{IE}}$ be the domain of parameter

H	Δ	J	$\Delta \cdot I$	K	F	Type of conic
> 0	$\neq 0$	> 0	> 0			Imaginary Ellipse
> 0	0	0		> 0		imaginary parallel lines
0				0	$\neq 0$	Poles

Table 4.4: Imaginary objects and Poles

vectors for imaginary ellipses. It follows from classification table 4.4 that

(4.15)

$$\mathbb{D}_{\text{IE}} = f_{\Delta}^{-1}((-\infty, 0) \bigcup (0, +\infty)) \bigcap f_{J}^{-1}((0, +\infty)) \bigcap f_{\Delta I}^{-1}((0, +\infty)) \bigcap f_{H}^{-1}((0, +\infty))$$

Thus $\mathbb{D}_{\text{IE}}$ is an open set on $\mathbb{S}^5$.

The domain $\mathbb{D}_{\text{IPL}}$ that represents vectors for imaginary parallel lines is neither open nor closed, which could be easily followed using the similar way of checking openness and closedness used for degenerate conics.

The parameter vector $(0, 0, 0, 0, 0, 1)$ and $(0, 0, 0, 0, 0, -1)$ are recognized as a empty solution which does not belong to any imaginary type. They occupy a domain of two points on $\mathbb{S}^5$, which is a closed set.

Imaginary objects can not be considered as a valid solution for our model and thus we include them into the empty set $\mathbb{D}_0$. Take a sequence of points corresponding to imaginary ellipses:

$$\mathbf{P}_n = (\frac{2n}{\sqrt{5n^2+1}}, 0, \frac{n}{\sqrt{5n^2+1}}, 0, 0, \frac{1}{\sqrt{5n^2+1}}) \rightarrow (\frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}, 0, 0, 0)$$

where $(\frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}, 0, 0, 0)$ corresponding to the point $(0, 0)$. A sequence in $\mathbb{D}_0$ converges to a limit outside of $\mathbb{D}_0$. So $\mathbb{D}_0$ is not closed.

Also notice that every neighborhood of point $(0, 0, 0, 0, 0, 1)$ corresponding to "empty set" contains a parameter vector $(0, 0, 0, \varepsilon, 0, \sqrt{1-\varepsilon})$ corresponding to a single line. Therefore $\mathbb{D}_0$ is not open either.

Compactness The unit sphere is a bounded and closed set in R^6. Thus it is compact where any continuous function should have global minimum and maximum.

Type of conic	Percentage
Ellipse	20.7
Hyperbola	76.8
Imaginary ellipse	2.5
Others	0

Table 4.5: Estimated volume of each open domain using adopted algebraic parameter

However, since a conic can be arbitrarily far away the any given set of points, objective function that represents the sum of squares of distances has no maximum. So either the objective function is not continuous or the space is not compact? We will prove that the objective function is continuous later in the next section Continuity of the objective function (defined on $\mathbb{S}^5$). So it is reasonable to blame on the non-compactness of the space. In fact, the absence of maximum is caused by the non-compactness of the subspace which consists of points corresponding to valid solution for our model and makes the complement of the space of empty solutions on the unit sphere. It is quite obvious that one can not evaluate the objective function from a given point to an imaginary object. So the objective function is only defined on the subspace containing ellipses,hyperbola,parabola and degenerate conics. In the previous discussion, we see that the space of empty sets is neither closed nor open. Its complement can not be a closed set. Therefore the subspace containing valid solutions is not compact.

4.4. Volumes of domains of conics

Here we address an interesting issue: "volume" of each domain on $\mathbb{S}^5$.

By running a Monte-Carlo simulation for which we picked $10,000$ random points from S^5, we find the percentage of each type of conics including the imaginary objects(see table 4.5)

Type of conic	Percentage
Ellipse	26.5
Hyperbola	65
Imaginary ellipse	8.5
Others	0

Table 4.6: Estimated volume of each domain using standard algebraic parameter

The table shows that hyperbola, ellipse and imaginary ellipse dominate the unit sphere. Any other conics are in minority classes. In the last section, we proved that hyperbola, ellipse and imaginary ellipse have open domains on the unit sphere. So they have nonempty interior and therefore positive measure while the others are neither open nor closed and they have empty interior, implying zero measure.

Adjusted algebraic parameter scheme In the section 4.1, we introduced two different algebraic parameters for conics: The first one (known as algebraic parameter) is widely recognized by community.

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

The second one that we call adjusted algebraic parameter appears to be almost the same as algebraic parameter except that the B, D and E are combined with the coefficient 2.

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

Here we explain the advantage the second parameter scheme takes over the first one. By imposing constraint $A^2 + \cdots + F^2 = 1$, we run Monte-Carlo simulation to estimate the percentage of each domain of conics under the algebraic parameter scheme (see tabel 4.6)

Under such parameter scheme, we have a much higher percentage of imaginary ellipse among randomly generated points. So by adding coefficient 2, we can expand the domain of points that represents valid solutions for our model. When running a

iterative numerical algorithm searching for the minimum of the objective, one should avoid a situation that the iteration traps into the domain of empty set and return a empty solution. The adjusted algebraic parameters resolve such a problem in a simple way: one only needs to multiply B,D,E from the standard algebraic parameters by 2 respectively and reduce chance of achieving empty solutions.

4.5. Boundaries of open domains

Reminder: the partition of $\mathbb{S}^5$ into domains

Recall that the unit sphere $\mathbb{S}^5$ (the parameter space) is divided into 10 domains corresponding to the main conic types, plus two extra points, the poles:

$$\mathbb{S}^5 = \mathbb{D}_{\rm E} \cup \mathbb{D}_{\rm H} \cup \mathbb{D}_{\rm IE} \cup \mathbb{D}_{\rm P} \cup \mathbb{D}_{\rm SP} \cup \mathbb{D}_{\rm IL} \cup \mathbb{D}_{\rm PL} \cup \mathbb{D}_{\rm IPL} \cup \mathbb{D}_{\rm CL} \cup \mathbb{D}_{\rm SL} \cup \{\mathbf{P}_1\} \cup \{\mathbf{P}_{-1}\}.$$

The subscripts are the codes of the conic types: E - ellipses, H- hyperbolas, IE - imaginary ellipses, etc.

The domains $\mathbb{D}_{\rm E}$, $\mathbb{D}_{\rm H}$, and $\mathbb{D}_{\rm IE}$ are five-dimensional, open, and they cover 100% of the sphere $\mathbb{S}^5$, in terms of volume (numerical experiment by H.M). The other domains have lower dimensionality and zero volume. They, in a sense, make pieces of the boundaries of the principal domains $\mathbb{D}_{\rm E}$, $\mathbb{D}_{\rm H}$, and $\mathbb{D}_{\rm IE}$.

Four-dimensional domains (hypersurfaces)

Four-dimensional domains $\mathbb{D}_{\rm P}$ (parabolas), $\mathbb{D}_{\rm SP}$ (single points), and $\mathbb{D}_{\rm IL}$ (intersecting lines)in a five-dimensional space play a prominent role: they separate the space into two parts (at least, locally). In geometry, they are called "hypersurfaces". Our hypersurfaces $\mathbb{D}_{\rm P}$, $\mathbb{D}_{\rm SP}$, and $\mathbb{D}_{\rm IL}$ separate our open domains $\mathbb{D}_{\rm E}$, $\mathbb{D}_{\rm H}$, and $\mathbb{D}_{\rm IE}$, or their components, from each other; see below.

Single points $\mathbb{D}_{\rm SP}$

The hypersurface $\mathbb{D}_{\rm SP}$ separates the open domain $\mathbb{D}_{\rm E}$ of ellipses from the open domain $\mathbb{D}_{\rm IE}$ of imaginary ellipses. To illustrate this fact, consider the parameter vector $\mathbf{P}_c = (-1, 0, -1, 0, 0, c)$, where c will play the role of a small variable (we will not normalize $\mathbf{P}_c$ to keep our formulas simple). This parameter vector corresponds

to the quadratic function

$$Q(x, y) = -x^2 - y^2 + c.$$

For $c > 0$, the equation $Q(x, y) = 0$ defines a small ellipse (more precisely, a small circle of radius $\sqrt{c}$), i.e., $\mathbf{P}_c \in \mathbb{D}_{\mathrm{E}}$ for $c > 0$. For $c = 0$, it is a single point, $(0, 0)$, i.e., $\mathbf{P}_0 \in \mathbb{D}_{\mathrm{SP}}$. For $c < 0$ it is an imaginary ellipse, i.e., $\mathbf{P}_c \in \mathbb{D}_{\mathrm{IE}}$ for $c < 0$. As c changes from small positive values to zero and then on to small negative values, the ellipse shrinks and collapses to a single point, and then disappears altogether (transforms into an imaginary ellipse). In the parameter space $\mathbb{S}^5$, this process corresponds to a continuous motion from the domain $\mathbb{D}_{\mathrm{E}}$ to the domain $\mathbb{D}_{\mathrm{IE}}$, across the hypersurface $\mathbb{D}_{\mathrm{SP}}$.

Intersecting lines $\mathbb{D}_{\mathrm{IL}}$

The hypersurface $\mathbb{D}_{\mathrm{IL}}$ separates the two components $\mathbb{D}_{\mathrm{H}}^+$ and $\mathbb{D}_{\mathrm{H}}^-$ of the open domain $\mathbb{D}_{\mathrm{H}}$ of hyperbolas from each other. To illustrate this fact, consider the parameter vector $\mathbf{P}_c = (-1, 0, 1, 0, 0, c)$, where c will again play the role of a small variable. This parameter vector corresponds to the quadratic function

$$Q(x, y) = -x^2 + y^2 + c. \tag{4.16}$$

The equation $Q(x, y) = 0$ defines a hyperbola with center $(0, 0)$, unless $c = 0$, in which case it is a pair of intersecting lines, $y = \pm x$. More precisely, for $c > 0$ it is a hyperbola with a "positive center", because $Q(0, 0) > 0$, i.e., $\mathbf{P}_c \in \mathbb{D}_{\mathrm{H}}^+$ for $c > 0$. For $c < 0$, it is a hyperbola with a "negative center", because $Q(0, 0) < 0$, i.e., $\mathbf{P}_c \in \mathbb{D}_{\mathrm{H}}^-$ for $c < 0$. For $c = 0$, it is a pair of intersecting lines, i.e., $\mathbf{P}_0 \in \mathbb{D}_{\mathrm{IL}}$. As c changes from small positive values to zero and then on to small negative values, the hyperbola with a positive center transforms into a pair of interesting lines and then into a hyperbola with a negative center. In the parameter space, this process corresponds to a continuous motion from the subdomain $\mathbb{D}_{\mathrm{H}}^+$ to the subdomain $\mathbb{D}_{\mathrm{H}}^-$, across the hypersurface $\mathbb{D}_{\mathrm{IL}}$ (see Figure 4.1).

Parabolas $\mathbb{D}_{\mathrm{P}}$

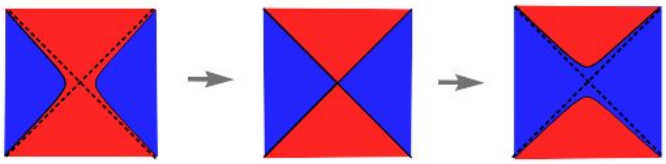


Figure 4.1: Red color corresponds to positive values of $Q(x, y)$ and blue color to its negative values.

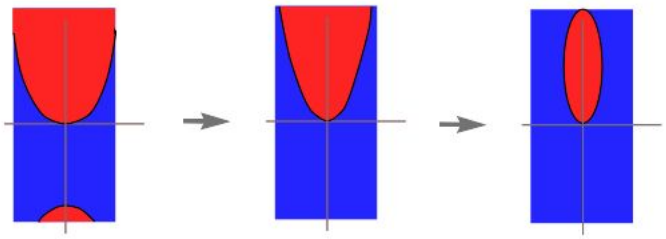


Figure 4.2: Red color corresponds to positive values of Q(x,y) and blue color to its negative values.

The hypersurface $\mathbb{D}_{\mathrm{P}}$ separates the domain $\mathbb{D}_{\mathrm{H}}$ of hyperbolas from the domain $\mathbb{D}_{\mathrm{E}}$ of ellipses. To illustrate this fact, consider the parameter vector $\mathbf{P}_c = (-1, 0, c, 0, 1, 0)$, where c will again play the role of a small variable. This parameter vector corresponds to the quadratic function

$$Q(x, y) = -x^2 + cy^2 + y = -x^2 + c\Big(y + \frac{1}{2c}\Big)^2 - \frac{1}{4c}.$$

For $c > 0$, the equation $Q(x, y) = 0$ defines a hyperbola, i.e., $\mathbf{P}_c \in \mathbb{D}_{\mathrm{H}}$ for $c > 0$. For $c = 0$, it is a parabola $y = x^2$, i.e., $\mathbf{P}_0 \in \mathbb{D}_{\mathrm{P}}$. For $c < 0$, it is an ellipse, i.e., $\mathbf{P}_c \in \mathbb{D}_{\mathrm{E}}$ for $c < 0$. As c changes from small positive values to zero and then on to small negative values, the hyperbola transforms into a parabola, and then into an ellipse. In the parameter space, this process corresponds to a continuous motion from the domain $\mathbb{D}_{\mathrm{H}}$ to the domain $\mathbb{D}_{\mathrm{E}}$, across the hypersurface $\mathbb{D}_{\mathrm{P}}$ (see Figure 4.2).

Sign changes

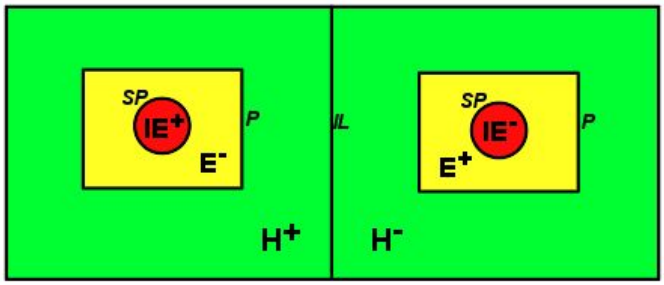


Figure 4.3: Principal domains and separating hypersurfaces. The labels correspond to our notation in the text: $\mathbb{H}^+$ means $\mathbb{D}^H_+$, etc.

A closer look at the above examples reveals that the "positive" subdomain $\mathbb{D}^+_{\mathrm{H}}$ borders on the "negative" subdomain $\mathbb{D}^-_{\mathrm{E}}$, and vice versa. Similarly, the "positive" subdomain $\mathbb{D}^+_{\mathrm{E}}$ borders on the "negative" subdomain $\mathbb{D}^-_{\mathrm{IE}}$, and vice versa. Thus the "sign" always changes when a parametr vector moves continuously from one subdomain to another.

A simplistic diagram

The above analysis is summarized in the following schematic diagram illustrating the structure of the parameter space, with all principal subdomains and the respective separating hypersurfaces (see Figure 4.3).

4.6. Fine structure of parameter space

In the previous section, we only described the main boundaries of the open domains, which are the hypersurfaces $\mathbb{D}_P$ (parabolas) and $\mathbb{D}_{IL}$ (intersecting lines), and $\mathbb{D}_{SP}$ (single points). We did not include domains of smaller dimension, i.e., $\mathbb{D}_{PL}$ (parallel lines) and $\mathbb{D}_{IPL}$ (imaginary parallel lines) of dimension three, as well as $\mathbb{D}_{CL}$ (coincident lines) and $\mathbb{D}_{SL}$ (single lines) of dimension two. Here we present a bigger picture that includes all of these elements.

Smaller boundary pieces Generally, domains of higher dimension terminate on domains of smaller dimension. In other words, domains of smaller dimension make boundaries of domains of higher dimension. More precisely, we say that a domain $\mathbb{D}_1$

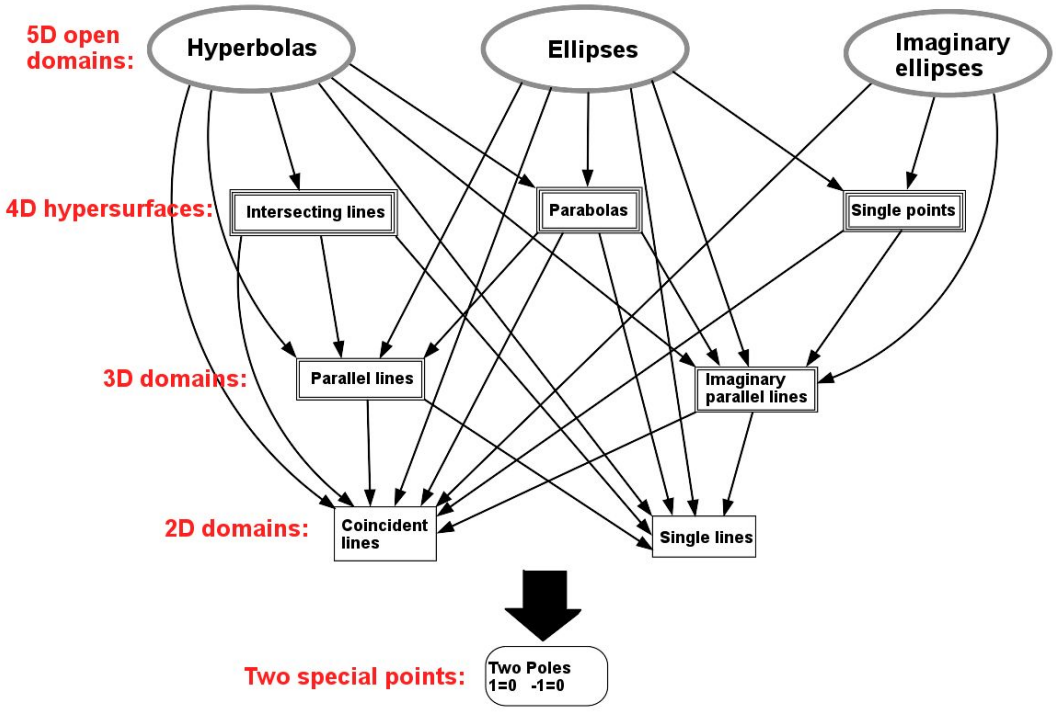


Figure 4.4: Boundary structure of domains

terminates on a domain $\mathbb{D}_2$ if there is a sequence of points $\mathbf{P}_n \in \mathbb{D}_1$ that converges to a point $\mathbf{P} \in \mathbb{D}_2$, i.e., $\mathbf{P}_n \to P$ as $n \to \infty$.

The diagram 4.4 above shows how our domains terminate on each other. An arrow from $\mathbb{D}_1$ to $\mathbb{D}_2$ means that $\mathbb{D}_1$ terminates on $\mathbb{D}_2$, i.e., there is a sequence of points of $\mathbb{D}_1$ that converges to a point of $\mathbb{D}_2$. The domains are named by the types of conics. The diagram contains all parts of the parameter space: from the largest, five-dimensional open domains to the smallest, two-dimensional regions. We note that all our domains terminate on each pole, so there should be an arrow from every domain down to the bottom line (Two Poles). For simplicity, we just put one large arrow pointing to the poles.

Two types of convergence The diagram 4.4 shows how parameter vectors $\mathbf{P} \in \mathbb{S}^5$ may converge, from one domain to another. Since this convergence involves algebraic parameters $\mathbf{P} = (A, B, C, D, E, F)$, we will call it "algebraic convergence".

It will refer to the convergence of a sequence of parameter vectors $\mathbf{P}_n$ to a parameter vector $\mathbf{P}$ on the sphere $\mathbb{S}^5$.

On the other hand, a sequence of conics S_n may converge to a conic S, in the sense of section (2.2). We will call this "geometric convergence". In particular, a sequence of conics of one type may geometrically converge to a conic of another type. In section (2.8) we described several examples of such a convergence: circles converging to a line, ellipses converging to a parabola, etc.

A natural question is: **Do algebraic and geometric types of convergence agree?**

They answer is **YES**, algebraic convergence agrees with geometric convergence in most cases, but there are some notable exceptions.

THEOREM 4.1. *(Convergence of conics: general case) Suppose a sequence $\mathbf{P}_i$ of parameter vectors corresponding to real (not imaginary) conics, S_n, converges to a parameter vector $\mathbf{P}$ corresponding to a real (not imaginary) conic, S, which is not a pair of coincident lines, i.e., $\mathbf{P} \notin \mathbb{D}_{\mathrm{CL}}$. Then $S_n \to S$ geometrically, in the sense of section* (2.2).

THEOREM 4.2. *(Divergence of conics: general case) Suppose a sequence $\mathbf{P}_n$ of parameter vectors corresponding to real (not imaginary) conics, S_n, converges to a parameter vector $\mathbf{P}$ corresponding to an imaginary conic or to a pole, i.e., $\mathbf{P} \in \mathbb{D}_{\mathrm{IPL}}$ or $\mathbf{P} = \mathbf{P}_{\pm 1}$. Then S_n moves off toward infinity, i.e., for any point $P = (x, y) \in \mathbb{R}^2$ we have dist$(P, S_n) \to \infty$ as $n \to \infty$.*

We note that the limit vector $\mathbf{P}$ cannot be in the domain of imaginary ellipses $\mathbb{D}_{\mathrm{IE}}$, because the latter is open.

THEOREM 4.3. *(The exceptional case of coincident lines) Suppose a sequence $\mathbf{P}_n$ of parameter vectors corresponding to real (not imaginary) conics, S_n, converges to a parameter vector $\mathbf{P} \in \mathbb{D}_{\mathrm{CL}}$ corresponding to a pair of coincident lines; the latter make a line in $\mathbb{R}^2$ which we denote by L. Then S_n gets closer and closer to L, as n grows.*

More precisely, for any rectangle

$$R = \{-A \le x \le A, \quad -B \le y \le B\}$$

we have

$$\max_{P \in S_n \cap R} \ \text{dist}(P, L) \to 0 \qquad \text{as} \quad n \to \infty.$$

In other words, if we look "through the window" R, we will see that all the points of S_n get closer and closer to L.

The proof of Theorem (4.1)-(4.3) are given in section (4.6).

Examples of algebraic convergence to coincident lines

On the other hand, the conics S_n in Theorem (4.3) may <u>not</u> converge to the line L in the sense of section (2.2). For example, S_n may be parabolas that converge to a half-line that is only a part of L. Or S_n may be hyperbolas that converge to two opposite half-lines that are only parts of L (see illustrations in section section (2.8)). Or S_n may be ellipses that converge to a segment of L (see illustrations in section section (2.8)). Or S_n may be single points that converge to a point in L. Or S_n may be any of the above but instead of converging to any part of L they may wander along L back and forth, or go off toward infinity.

For example, let S_n be defined by

$$x^2 + \frac{\alpha_n (y + C_n)^2}{1 + C_n^2} = \beta_n,$$

where $\alpha_n \to 0$ and $\beta_n \to 0$ as $n \to \infty$. Then algebraically this sequence converges to $x^2 = 0$, which is a pair of coincident lines. But geometrically S_n may be an ellipse or a hyperbola or a single point, depending on the values (and the signs) of C_n, α_n, β_n, and it may converge to various parts of L or move back and forth along L or move off toward infinity altogether.

CHAPTER 5

OBJECTIVE FUNCTION FOR QUADRATIC CURVES (CONICS)

In this chapter we study the objective function on the parameter space of conics, i.e., on the sphere $\mathbb{S}^5$. Recall that given some points $P_1, \ldots, P_n \in \mathbb{R}^2$, the objective function is the sum of squares of the distances to a model object (in our case, conic) S:

$$\mathcal{F}(S) = \sum_{i=1}^{n} \big[\operatorname{dist}(P_i, S)\big]^2, \tag{5.1}$$

The objective function depends on the points $P_1, \ldots, P_n$, but in practical settings those are fixed, so the only variable is S, which is regarded as the sole argument of $\mathcal{F}$.

5.1. Continuity of the objective function on the sphere

Domain of the objective function The objective function (5.1) can be transformed into an equivalent form with parameter vector as its argument:

$$\mathcal{F}(\mathbf{P}) = \sum_{i=1}^{n} \big[\operatorname{dist}(P_i, \mathbf{P})\big]^2, \tag{5.2}$$

Now $\mathcal{F}$ becomes a function defined on $\mathbb{S}^5$, except domains corresponding to imaginary conics or poles:

$$\mathfrak{D}_{\mathcal{F}} = \mathbb{D}_{\mathrm{E}} \cup \mathbb{D}_{\mathrm{H}} \cup \mathbb{D}_{\mathrm{P}} \cup \mathbb{D}_{\mathrm{SP}} \cup \mathbb{D}_{\mathrm{IL}} \cup \mathbb{D}_{\mathrm{PL}} \cup \mathbb{D}_{\mathrm{CL}} \cup \mathbb{D}_{\mathrm{SL}}.$$

We note that $\mathfrak{D}_{\mathcal{F}}$ does not include regions $\mathbb{D}_{\mathrm{IE}}$ and D_{IPL} corresponding to imaginary conics, and it does not include the poles $\mathbf{P}_{\pm 1}$.

THEOREM 5.1. ***Continuity of the objective function*** *The objective function $\mathcal{F}$ is continuous everywhere on its domain $\mathfrak{D}_{\mathcal{F}}$ except on the region $\mathbb{D}_{\mathrm{CL}}$ corresponding to coincident lines.*

This theorem is an immediate consequence of Theorem 4.1 of section 4.6 and main theorem of section 2.4. Indeed, we only need to apply a general principle: the composition of two continuous functions is a continuous function.

Lower semi-continuity of the objective function

On the region $\mathbb{D}_{\mathrm{CL}}$ corresponding to coincident lines the objective function $\mathcal{F}$ is badly discontinuous, according to Theorem 4.3 (and the discussion around it) in section 4.6. We showed there that if $\mathbf{P}_n \to \mathbf{P}$ and $\mathbf{P} \in \mathbb{D}_{\mathrm{CL}}$ corresponds to a coincident line L, then the conics S_n corresponding to $\mathbf{P}_n$ may move back and forth along the line L or move off toward infinity. Accordingly, the values of the objective function $\mathcal{F}(\mathbf{P}_n)$ may oscillate within a wide range or diverge to infinity.

But the objects S_n must get closer and closer to L, as n grows, they just may not stretch all the way along L. This implies that the objects S_n, in the limit $n \to \infty$, cannot provide a better fit to the given points than the line L does. Thus any limit value obtained from $\mathcal{F}(\mathbf{P}_n)$ cannot be smaller than the value $\mathcal{F}(\mathbf{P})$, i.e.,

$$\liminf_{n\to\infty} \quad \mathcal{F}(\mathbf{P}_n) \geq \mathcal{F}(\mathbf{P}).$$

Hence we obtain one more important fact:

THEOREM 5.2. ***Lower semi-continuity of the objective function*** *The objective function $\mathcal{F}$ is lower semi-continuous on the region $\mathbb{D}_{\mathrm{CL}}$ corresponding to coincident lines.*

Remember a function $f(x)$ is lower semi-continuous at x_0 if

$$\liminf_{x\to x_0} f(x) \geqslant f(x_0)$$

Since it is already known from Theorem (5.1) that $\mathcal{F}(\mathbf{P})$ is continuous everywhere except on the region $\mathbb{D}_{\mathrm{CL}}$, it remains to show that $\mathcal{F}(\mathbf{P})$ is lower semi-continuous

within $\mathbb{D}_{\mathrm{CL}}$. Furthermore, it is enough to verify that the distance function $\mathrm{dist}(P_i, S)$ is lower semi continuous at every point p within the domain of coincident lines. We will prove for any given point $P \in R^2$

$$\liminf_{\mathbf{P}_i \to \mathbf{P}_0} \mathrm{dist}(P, \mathcal{G}(\mathbf{P}_i)) \geqslant \mathrm{dist}(P, \mathcal{G}(\mathbf{P}_0))$$

where $\mathcal{G}(\mathbf{P})$ is the conic represented by $\mathbf{P}_i$.

Proof. Let $\mathbf{P}_i = (A_i, B_i, C_i, D_i, E_i, F_i)$ be a sequence of parameter points with a limit $\mathbf{P}_0$ corresponding to a pair of coincident lines. Since $\mathrm{dist}(P, \mathcal{G}(\mathbf{P}_i)) = \inf \mathrm{dist}(P, P_i)$ where $P_i = (x_i, y_i) \in \mathcal{G}(\mathbf{P}_i)$,

$$Q(x_i, y_i | p_i) = A_i x_i^2 + 2B_i x_i y_i + C_i y_i^2 + 2D_i x_i + 2E_i y_i + F_i = 0.$$

Let us consider the following cases:

(i) $\liminf_{i \to \infty} \mathrm{dist}(P, P_i) = c$ $(c \geqslant 0)$.

There exists a subsequence of $\{P_i\}$, denoted by $\{P_{i_j}\}$ such that

$$\lim_{i_j \to \infty} \mathrm{dist}(P, P_{i_j}) = c.$$

Furthermore, since $\{P_{i_j}\}$ is bounded, one can always find a convergent subsequence with a limit $P_0 = (x_0, y_0)$. For simplicity, let us assume $\{P_{i_j}\}$ is convergent so that

$$\lim_{i_j \to \infty} \mathrm{dist}(P, P_{i_j}) = \mathrm{dist}(P, P_0) = c$$

Then

$$\lim_{i_j \to \infty} Q(x_{i_j}, y_{i_j} | \mathbf{P}_{i_j}) = Q(x_0, y_0 | \mathbf{P}_0) = 0.$$

The last equality indicates $P_0 = (x_0, y_0) \in \mathcal{G}(\mathbf{P}_0)$. So

$$\mathrm{dist}(P, \mathcal{G}(p_0)) \leqslant \mathrm{dist}(P, P_0) \leqslant \liminf_{i \to \infty} \mathrm{dist}(P, P_i) = c$$

(ii) $\liminf_{i \to \infty} \mathrm{dist}(P, P_i) = \infty$

It is quite obvious that $\mathrm{dist}(P, \mathcal{G}(\mathbf{P}_0)) < \infty$. Therefore,

$$\mathrm{dist}(P, \mathcal{G}(\mathbf{P}_0)) < \liminf_{i \to \infty} \mathrm{dist}(P, \mathcal{G}(\mathbf{P}_i)).$$

The proof is completed and now we see that $\mathcal{F}'(\mathbf{P})$ is lower semi continuity within the domain of all real conics. □

THEOREM 5.3. ***Growth near the imaginary conics and poles*** *The objective function $\mathcal{F}$ grows to infinity near the region $\mathbb{D}_{\mathrm{IPL}}$ and near the poles $\mathbf{P}_{\pm 1}$. More precisely, if $\mathbf{P}_n \to \mathbf{P}$ and either $\mathbf{P} \in \mathbb{D}_{\mathrm{IPL}}$ or $\mathbf{P} = \mathbf{P}_{\pm 1}$, then $\mathcal{F}(\mathbf{P}_n) \to \infty$.*

This theorem follows immediately from Theorem of section (4.6).

The existence of a global minimum revisited It follows from Theorem (5.3) that the $\mathfrak{D}_{\mathcal{F}}$ is not compact. But we can cut out and ignore a small open vicinity of the region $\mathbb{D}_{\mathrm{IPL}}$ and the poles $\mathbf{P}_{\pm 1}$ where the function is too big. Then the remaining part of the domain $\mathfrak{D}_{\mathcal{F}}$ will be compact. And now the lower semi-continuity of $\mathcal{F}$ (proven in Theorem 5.1 + Theorem 5.1) guarantees the existence of its global minimum. Indeed, any lower semi-continuous function on a compact domain attains its minimum;

The existence of a global minimum is nothing new, however, as we have proved the existence of the best fitting object already in section (2).

We continue in next section.

5.2. Differentiability of the objective function on the sphere

Recall that the objective function is the sum of squares of the distances from the given (fixed) points $P_1, \ldots, P_n \in \mathbb{R}^2$ to a (variable) conic S:

$$\mathcal{F}(S) = \sum_{i=1}^{n} \big[\mathrm{dist}(P_i, S)\big]^2. \tag{5.3}$$

Its domain $\mathfrak{D}_{\mathcal{F}} \subset \mathbb{S}^5$ is a part of the sphere $\mathbb{S}^5$. It can be decomposed into subdomains corresponding to different conic types:

$$\mathfrak{D}_{\mathcal{F}} = \mathbb{D}_{\mathrm{E}} \cup \mathbb{D}_{\mathrm{H}} \cup \mathbb{D}_{\mathrm{P}} \cup \mathbb{D}_{\mathrm{SP}} \cup \mathbb{D}_{\mathrm{IL}} \cup \mathbb{D}_{\mathrm{PL}} \cup \mathbb{D}_{\mathrm{CL}} \cup \mathbb{D}_{\mathrm{SL}}.$$

In the previous section we showed that $\mathcal{F}$ is continuous on its entire domain $\mathfrak{D}_{\mathcal{F}}$ except a tiny subregion $\mathbb{D}_{\mathrm{CL}} \subset \mathfrak{D}_{\mathcal{F}}$ corresponding to coincident lines. On that latter region $\mathcal{F}$

is badly discontinuous, but at least we established that it was lower semi-continuous. In this section we investigate the differentiability of $\mathcal{F}$.

Differentiability of objective function: General considerations The derivatives of given function $\mathcal{F}$ is often in fitting algorithms (such as the steepest descent, Newton-Raphson, Gauss-Newton, or Levenberg-Marquardt). Some use the first order derivative of $\mathcal{F}$, others use the second order derivative, or approximations to the second order derivative.

Thus it is essential that our objective function $\mathcal{F}$ is differentiable, at least once. As $\mathcal{F}$ is the sum of squares of the distances, see (5.1), it will be enough to check that $[\text{dist}(P_i, S)]^2$, i.e., the square of the distance from the given point $P_i = (x_i, y_i) \in \mathbb{R}^2$ to the conic S, is differentiable with respect to the conic's parameters.

We consider a more general problem. Given a point $P = (x_0, y_0)$ and a conic S, we will investigate the differentiability (with respect to the parameters of S) of the function

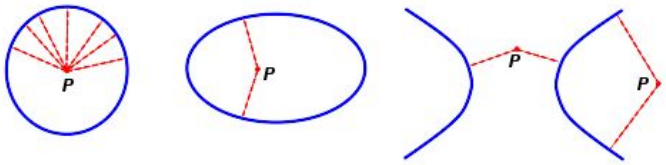
$$[\text{dist}(P, S)]^2 = [\text{dist}(P, Q)]^2 = (x - x_0)^2 + (y - y_0)^2,$$

where $Q = (x, y)$ is the projection of P onto the conic S; see section 2.1. To this end it will be enough to check that the coordinates x, y of the footpoint Q of the projection are differentiable with respect to the conic's parameters.

One may guess, intuitively, that whenever the point $P = (x_0, y_0)$ is kept fixed and the conic S changes continuously, the projection Q of P onto S would change continuously and smoothly. We will prove that generally this is true. However, there are exceptional cases where the continuity breaks down.

Differentiability of objective function: Some exceptions

The reason for the breakdown is that the point Q on the conic S closest to the given point P may be not unique. For example, if S is a circle and P is its center, then all the points of S are equally distant from P, hence the point Q can be chosen anywhere on the circle. Another example: S is an ellipse and P lies on the major axis near the center. Then there are exactly two points on S closest to P (they



are symmetric about the major axis of S). Similar situations occur when S is a hyperbola or a parabola and P lies on its axis. See illustrations in the figure 5.2. In these exceptional cases, if one changes such a conic S continuously, then the point Q may instantaneously "jump" from one side (or branch) of S to another.

A more subtle exceptional case occurs when P lies at the center of curvature of S at the point Q. This means that P coincides with the center of the osculating circle [**?**] of the conic S at the projection point Q. Then the projection Q may be technically unique, but "barely unique", as to the second order all the points on S close enough to Q will be equally distant from P. This is a subtle situation, we will explore it separately.

THEOREM 5.4. *(**Differentiability of projection coordinates**) Let S be a conic and P a given point. Suppose (i) the point Q on the conic S closest to the given point P is unique and (ii) P is not the center of curvature of the conic S at the point Q. Then the coordinates x and y of the point Q are differentiable with respect to the conic's parameters.*

Theorem 5.4 is proved by implicit differentiation. See the proof in section A.5.

Our proof ensures the existence of the first order derivatives of x and y with respect to the conic's parameters. One can easily extend it to higher order derivatives, so that x and y can be shown to have derivatives of all orders.

We note, however, that for practical purposes first order derivatives suffice. The most popular minimization algorithms, such as Gauss-Newton and Levenberg-Marquardt [**23, 25**], use approximations to the second order derivatives of the objective function $\mathcal{F}$, and those approximations only require the first derivatives of the distances

dist(P_i, S)'s. That is, only the first order derivatives of the coordinates of projection points x and y with respect to the conic's parameters are needed.

Next we turn to the exceptional case (ii) in the above theorem, i.e., suppose P is at the center of the osculating circle of S at the point Q. Then the coordinates x and y are not differentiable with respect to the conic's parameters. But surprisingly the distance dist(P, S) is differentiable with respect to the conic's parameters.

Example of a point at center of osculating circle As a simple example, consider a family of parabolas defined by equation

$$y = x^2 - C,$$

where C is the only scalar parameter. For $C = 0$ we have the classical parabola $y = x^2$. At the point $Q = (0, 0)$ of this parabola, the osculating circle has center $(0, 1/2)$ and radius $1/2$. So let us put the data point $P = (0, 1/2)$ right at the center of the osculating circle. Then for $C = 0$ we get the (unique) projection $Q = (0, 0)$ of the point P onto the parabola $y = x^2$ and the distance from P to the parabola is $d(0) = 1/2$.

Let us see what happens when C becomes positive or negative. For $C < 0$, the parabola moves up, toward the point P. That point still has a unique projection $Q = (0, -C)$ onto the parabola $y = x^2 - C$. Thus the coordinates of the projection point $x(C) = 0$ and $y(C) = -C$ appear to be smooth functions.

However, for $C > 0$, the parabola moves down, away from the point P. Now P has two projections onto the parabola $y = x^2 - C$; their footpoints are $Q_\pm = (\pm\sqrt{C}, 0)$. Thus the coordinates of the projection points are $x(C) = \pm\sqrt{C}$ and $y(C) = 0$.

We see that the y-coordinate $y(C)$ of the projection point changes its slope at $C = 0$, from $y' = -1$ for $C < 0$ to $y' = 0$ for $C > 0$. Thus the function $y(C)$ is not differentiable at $C = 0$, even though it has one-sided derivatives. The x-coordinate $x(C)$ of the projection point does not even have a one-sided derivative corresponding to $C > 0$. Indeed, $x'(C) = \pm 1/(2\sqrt{C})$, which approaches infinity as $C \to 0$. Hence the slope of the function $x(C)$ at $C = 0$ is vertical, its derivative turns infinite!

We see that both coordinates, $x(C)$ and $y(C)$, of the projection point fail to be differentiable at the parameter value $C = 0$, which corresponds to the data point P right at the center of the osculating circle.

However, the distance $d(C)$ from the fixed point $P = (0, 1/2)$ to the parabola $y = x^2$ miraculously remains smooth, even at $C = 0$! Indeed, by elementary calculations

$$d(C) = 1/2 + C \quad \text{for} \quad C < 0 \qquad \text{and} \qquad d(C) = \sqrt{\frac{1}{4} + C} \qquad \text{for} \quad C > 0.$$

Thus its derivative is

$$d'(C) = 1 \quad \text{for} \quad C < 0 \qquad \text{and} \qquad d'(C) = \frac{1}{2\sqrt{\frac{1}{4} + C}} \qquad \text{for} \quad C > 0.$$

We see that at the junction point $C = 0$ both derivatives coincide, they are equal to 1. Thus the function has a continuous first order derivative at the value $C = 0$.

Remark: We must note, however, that the second order derivative $d''(C)$ is <u>not</u> continuous at $C = 0$, thus it does <u>not</u> exist at this point.

Still, the existence and continuity of the first derivative looks like a stunning miracle here. It calls for an explanation.

Fortunately, the existence and continuity of the first derivative is not just a sheer luck that we observed in one particular example. It is a general fact that we prove in section A.5 (see **smoothness at centers of osculating circles**).

Thus we get

THEOREM 5.5. *(**Differentiability of distances**) Let S be a conic and P a given point. Suppose (i) the point Q on the conic S closest to the given point P is <u>unique</u> and (ii) P coincides with the center of curvature of the conic S at the point Q. Then the distance* $\operatorname{dist}(P, S)$ *is differentiable with respect to the conic's parameters.*

The proof is given in section (A.5).

Thus the objective function $\mathcal{F}$ is differentiable, unless the point P has more than one projection onto the conic S.

Examples of non-differentiability

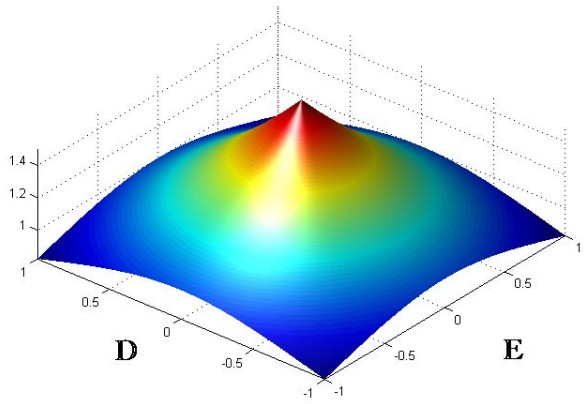


Figure 5.1: Example of Non-differntiability

In rare cases where the condition (i) of Theorems 1 and 2 does not hold, the objective function may not be differentiable. This happens, for instance, if S is an ellipse and one of the data points P_i happens to lie on its major axis somewhere in the middle of S (then P_i is equally distant from the two halves of the ellipse). Or if S is a circle and one of the data points P_i is its center.

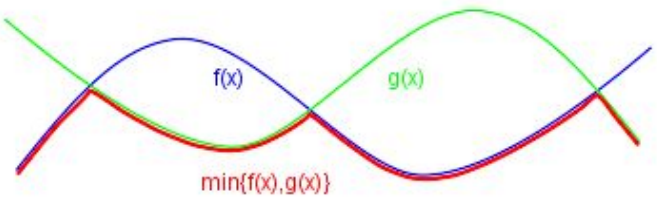
To illustrate the above effect let us consider a simplified family of conics defined by

$$x^2 + y^2 + 2Dx + 2Ey - 3 = 0$$

where only two algebraic parameters, D and E, are variable and all the others are fixed ($A = C = 1$, $B = 0$, and $F = -3$). This is actually a family of circles with center $(-D, -E)$ and radius $R = \sqrt{D^2 + E^2 + 3}$. The distance from $P = (x_0, y_0)$ to this circle is given by

$$\text{dist}(P, S) = \left| \sqrt{(D + x_0)^2 + (E + y_0)^2} - \sqrt{D^2 + E^2 + 3} \right|.$$

Figure 5.2is the graph of this distance, as a function of D and E, plotted by MATLAB . We have set $x_0 = y_0 = 0$ and let D and E vary from -1 to 1. We clearly see a sharp peak on the graph at the point $D = E = 0$, exactly where the point $P = (0, 0)$



coincides with the center of the circle. The graph is a cone-shaped surface near the peak with no derivatives at the summit.

To summarize, the cases of non-differentiability of the objective function $\mathcal{F}$ are rare. They occur when one of the data points happens to be in an unusual place where the distance to the conic may be computed in more than one way. Such points "confuse" the objective function and cause the failure of its differentiability.

No local minima at singular points

It is important to explore what happens whenever the objective function $\mathcal{F}$ fails to be differentiable. It turns out, fortunately, that in all such cases the shape of $\mathcal{F}$ resembles a "peak" (pointing upward), as in the above illustration. It cannot have a shape of a "pothole" (pointing downward).

Indeed, suppose a data point P_i can be orthogonally projected onto the conic S in more than one way (meaning its projections on different parts or on different branches of S). Denote the footpoints of those projections by Q_i', Q_i'', etc. Then

$$\operatorname{dist}(P_i, S) = \min\{\operatorname{dist}(P_i, Q_i'), \operatorname{dist}(P_i, Q_i''),\ \text{etc.}\}$$

Thus the distance is obtained as the minimum several smooth functions. And here is a general fact: the minimum of several smooth functions can only have "peak-type" singularities, not "pothole-type" singularities; see a simple illustration in figure 5.2. In other words, $\mathcal{F}$ has "peaks", or local maxima, at singular points.

THEOREM 5.6. *(**Smoothness at local minima**) The objective function $\mathcal{F}$ is smooth at all its local minima. More precisely, the first order derivatives of $\mathcal{F}$, as well as those of the distances* $\operatorname{dist}(P_i, S)$*, exist and are continuous at all local minima.*

See a proof in section (A.5).

Since our main goal is minimization of $\mathcal{F}$, i.e., finding its (local) minima, the singularities of $\mathcal{F}$ will not really concern us, they will not be harmful. Standard minimization algorithms, such as Levenberg-Marquardt or Trust Region, are prohibited from moving in the "wrong direction" where the function $\mathcal{F}$ increases. They will only move if they find a smaller value of $\mathcal{F}$. This restriction forces them to move away from local maxima of $\mathcal{F}$, in particular away from singular points of $\mathcal{F}$.

If an algorithm converges to a limit, then $\mathcal{F}$ has a local minimum there, and by our Theorem 3 the function $\mathcal{F}$ and the distances $\operatorname{dist}(P_i, S)$ have continuous first order derivatives. Since the above mentioned algorithms only use the first order derivatives of the distances $\operatorname{dist}(P_i, S)$, they should be able to find the local minimum of $\mathcal{F}$ and converge quickly.

Summary

To summarize, we list all the domains where the minimization algorithms are likely to "maneuver" searching for the best fitting conic and where the best fit can be found: Ellipses $\mathbb{D}_{\mathrm{E}}$, hyperbolas $\mathbb{D}_{\mathrm{H}}$, parabolas $\mathbb{D}_{\mathrm{P}}$, intersecting lines $\mathbb{D}_{\mathrm{IL}}$ and parallel lines $\mathbb{D}_{\mathrm{PL}}$. We formalize this in the following statement:

THEOREM 5.7. *(**Essential domain**) For any set of data points $P_1, \dots, P_n$ the global minimum of the objective function $\mathcal{F}$ belongs to the union*

$$\mathfrak{D}_{\mathcal{F},\mathrm{ESS}} = \mathbb{D}_{\mathrm{E}} \cup \mathbb{D}_{\mathrm{H}} \cup \mathbb{D}_{\mathrm{P}} \cup \mathbb{D}_{\mathrm{IL}} \cup \mathbb{D}_{\mathrm{PL}}. \tag{5.4}$$

If the objective function $\mathcal{F}$ has multiple global minima, then at least one of them belongs to the above union. This union cannot be shortened, i.e., for any conic S in this union of domains there exists a data set for which S provides the unique best fit.

See a proof in section (A.6).

We call $\mathfrak{D}_{\mathcal{F},\mathrm{ESS}}$ the "essential domain, or the essential part of the domain $\mathfrak{D}_{\mathcal{F},}$ of the function $\mathcal{F}$. The above theorem basically says that all the other parts of the

parameter space $\mathbb{S}^5$ can be ignored for the purpose of minimization of the objective function. On those parts $\mathcal{F}$ is either not defined or tends to grow.

CHAPTER 6

GEOMETRIC FIT AND THE PROBLEM OF THE MOMENTS

In EIV problems, the given points on the plane are subject to random errors (usually normal r.v). Due to the complexity of curve fitting problems the exact distribution and moments of parameters are hard to determine. It was already known that the parameters in linear regression $y = \alpha + \beta x$ do not have finite absolute moments

$$\mathbb{E}(|\hat{\alpha}|) = \mathbb{E}(|\hat{\beta}|) = \infty \tag{6.1}$$

This fact was found by Anderson in 1976 [**6**]. Recently, Chernov discovered that the parameters in circular regression $(x-a)^2 + (x-b)^2 = R^2$ have infinite moments,too [**14**]

$$\mathbb{E}(|\hat{a}) = \mathbb{E}(|\hat{b}|) = \mathbb{E}(|\hat{R}|) = \infty \tag{6.2}$$

In this chapter, we will investigate elliptical regression – fitting ellipses to observed points whose both coordinates are measured with errors. We prove under any standard assumptions on the statistical distribution of errors that are commonly adopted in the literature, the estimate for the major axis, center coordinates have infinite moments. The minor axis have finite first moment but infinite second moment. Our discussion will follow the general strategy used in [**2**] and [**14**].

6.1. Geometric elliptical fit

In the problem of fitting curve to given points, geometric fit which minimizes the sum of squares of distances from points to the curve is considered as the most reliable fitting method. Let us recognize the following important fact about the geometric fit.

THEOREM 6.1. *In functional model, the maximum likelihood estimator of the primary parameters $\theta_1, \ldots, \theta_k$ is attained on the curve that minimizes the sum of squares of orthogonal distances to data points.*

see [**13**] It is commonly regarded as the best (most accurate and reliable) fitting method. However the corresponding parameter estimates often have a bizarre feature: they do not have finite moment. On the other hand, classical estimates minimizing vertical distances, even in the linear case [**6**], are known to have finite moments. But in practical application, the classical estimates are known to be much less accurate and have heavy bias in the estimates of some parameters. In other words, paradoxically, a better estimate has infinite moments. (Theoretically, its mean squared error is infinite and bias can not be measured) while a worse estimate has finite moments (so its bias and mean squared error are finite).

Existence revisited

Before we proceed to any formal investigation, let us recall the issue of existence arises in elliptic regression: there is a nonzero probability that the best fitting ellipse would not exist. Strictly speaking, if one fits a quadratic curve (a conic section) to observed points, then the best fitting conic may be (i) an ellipse or (ii) a hyperbola or (iii) a parabola or (iv) a straight line or (v) a pair of straight lines. Even though lines and parabolas occur with probability zero (thus they can be ignored), hyperbolas occur with a positive probability and have to be reckoned with. When the best fitting conic is a hyperbola, then the problem of fitting ellipses has no solution (see section A.1 in appendix). In that case for any ellipse one can find another ellipse that fits the given points better than the previous one (in the sense of a smaller sum of squares of distances). A sequence of such ellipses that approximate these given points progressively well will converge to a parabola [**28**]. In a numerical experiment of fitting five random points generated by a continuous distribution (normal or uniform), the ellipse turns up in 30% of the cases while the hyperbola occurs in 70%. This suggests that there is a significant chance that the best fitting ellipse does not exists at all.

Therefore our analysis has to be restricted to data sets where the best fitting ellipse does exist (i.e., where the best fitting conic is an ellipse, rather than anything else). The expectations of the geometric parameters in this chapter should be understood as conditional expectations (i.e.,the integral of the estimates of the parameters for the data sets for which the best fitting ellipse does exists).

6.2. General Strategy

In ellipse fitting problem one estimates two axes a, b, coordinates of the center (C_x, C_y) and the angle of tilt α (the angle between the major axis the horizontal axis). The angle α is commonly assumed to be between 0 and 2π. So its estimate $\hat{\alpha}$ should have finite moment. The following analysis applies to the remaining parameters.

The absolute first moment of parameter estimate $\hat{\theta}$

$$\mathbb{E}(|\hat{\theta}|) = \int_0^\infty \text{Prob}(|\hat{\theta}| > x)\, dx \tag{6.3}$$

is infinite if the distribution has a power-law tail $\text{Prob}(|\hat{\theta}| > x) \sim x^{-\gamma}$ as $x \to \infty$ with $\gamma \leq 1$. The reciprocal $\zeta = 1/\hat{\theta}$ then satisfies $\text{Prob}(|\zeta| < y) \sim y^{\gamma}$ with $\gamma \leq 1$. It is easy to see that ζ vanishes as $\hat{\theta}$ grows to infinity. Thus we only need to check that ζ has a positive density function which does not vanish at 0.

Suppose we can position n points $(x_1, y_1), \ldots, (x_n, y_n)$ so that the parameter estimate $\hat{\theta}$ will be infinite and hence $\zeta = 0$. Also note that $\hat{\theta}$ and ζ are continuous function of coordinates $(x_1, y_1), \ldots, (x_n, y_n)$. Next we fix all coordinates except only, say x_1 varies.

LEMMA 6.1. *Suppose that the derivative $|\partial\zeta/\partial x_1| \leq D$ for some $D > 0$. Then the conditional absolute moment of $\hat{\theta}$ given that the coordinates $y_1, x_2, y_2, \ldots, x_n, y_n$ (i.e., all but x_1) are fixed, is infinite, i.e., $\mathbb{E}(|\hat{\theta}| \mid y_1, x_2, y_2, \ldots, x_n, y_n) = \infty$.*

PROOF. Since the original joint distribution of all the coordinates $x_1; y_1; \ldots; x_n; y_n$ has a strictly positive density, the conditional distribution of x_1 (given that all the

other coordinates are fixed) also has a strictly positive density. And since $|\partial\zeta/\partial x_1| \leq D$, the conditional density of ζ

$$f_\zeta(z|y_1;\ldots;x_n;y_n) = |\frac{d(\zeta^{-1}(z))}{dz}| \cdot f_{x_1}(\zeta^{-1}(z)) = |\frac{\partial x_1}{\partial \zeta}\ |_{\zeta=z}\ | \cdot f_{x_1}(\zeta^{-1}(z)) > 0 \tag{6.4}$$

Hence, as we have seen, the conditional expectation of $\hat{\theta}$ is infinite. □

The argument holds true if every coordinates is slightly changed, i.e., $x_2 \in I_i$ $(i = 2,\ldots,n)$ and $y_i \in J_j$ $(j = 1,\ldots,n)$ for some I_i and J_j. Therefore the unconditional expectation $\mathbb{E}(|\hat{\theta}|) = \infty$.

Next we will construct such an example for which the moments of parameters for an ellipse are infinite.

6.3. Elliptic regression (for five points)

When the total number of data points is $n = 5$, the argument is relatively simple (see [**2**]). We have seen in the section 2.8 that two semi axes $a \geq b$ grows to infinity as the ellipse converge to a parabola. Taking the clue from this fact, we can choose five points $(-2,1),(2,1),(-1,0),(1,0)$ and $(0,-1/3)$. Then we fix four small squares $\mathbb{B}_i$ $(i = 1,\ldots,4)$ of size $2h^2 \times 2h^2$ centered at each of first four points and one small rectangle $\mathbb{B}_5$ of size $2h^2 \times 2h$ (h is a small number such as 10^{-9}) at the last point. Then there always exists an interpolating object S_{best} passing through five points (quadratic curve or line) (see Figure 6.1).

We choose one points (x_0, y_0) in the rectangle $\mathbb{B}_5$ and one points (x_i, y_i) from each square $\mathbb{B}_i (i = 1,2,3,4)$. Note that y_0 is allowed to vary within a interval of size $2h$ while all other coordinates are restricted to much smaller interval of size $2h^2 \ll 2h$. By simple geometry, the best fitting (interpolating) object is an ellipse when $y_0 = -1/3 + h$ or a hyperbola when $y_0 = -1/3 - h$. Let us fixed all coordinates except y_0. As y_0 varies from $-1/3+h$ to $-1/3-h$, the interpolating ellipse converges to a parabola and than becomes hyperbola. We will only consider the part of the interval $(-1/3 + h^*, -1/3 + h)$ $(-h < h^* < h)$. The major axis $\hat{a}$ of the ellipse reach at infinity when $y_0 = -1/3 + h^*$ and the corresponding interpolating object is

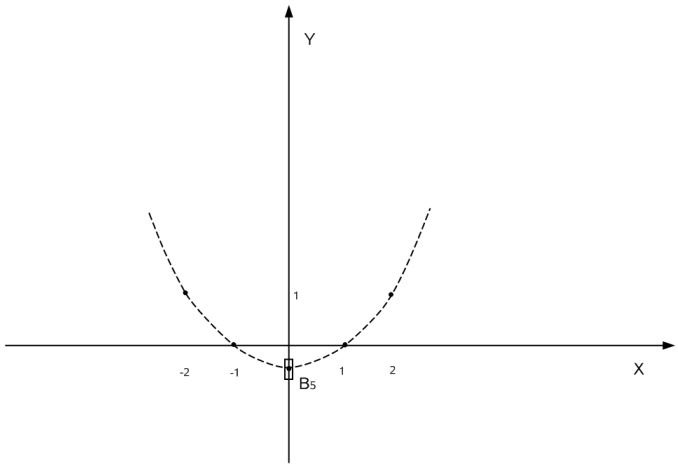


Figure 6.1: The best fitting conic for five points

a parabola. Let us define $\zeta = 1/\hat{a}$. Keep in mind that ζ is a function only depending on y_0. We will prove the following fact:

THEOREM 6.2. *For any fixed values $x_1, x_2, y_2, \ldots, x_5, y_5$, the function $\zeta(y_0)$ has bounded derivative, i.e. $|\partial\zeta(y_0)/\partial y_0| < D$ for some $D > 0$.*

The above fact implies our desired result $\mathbb{E}(|\hat{a}|) = \infty$. And since $\hat{C}_y \approx \hat{a} - 1/3$, $\mathbb{E}(|\hat{C}_y|) = \infty$. Rotating the construction by 2π, we can obtain $\mathbb{E}(|\hat{C}_x|) = \infty$. Besides, we will show that as the ellipse degenerates to parabola the minor axis $\hat{b}$ will grow to infinity as well but at a lower rate $\hat{b} \sim a^{1/2}$.

Proof of Theroem 6.2 Here we provide a proof for the infinite moment of some of parameters (major axis and coordinates of the center) for the best fitting ellipse which may illustrate our main idea of infinite moment for arbitrary $n > 5$. First we fix five points at $(-2, 1)$, $(2, 1)$, $(-1, 0)$, $(1, 0)$ and (x_0, y_0) where $x_0 \in (-2h^2, 2h^2)$ and $y_0 = z - 1/3 \in (-1/3 - 2h, -1/3 + 2h)(h = 10^{-9})$. Note that the best fitting (interpolating) curve (including parabolas) will be an ellipse when $x_0^2 \leq 3y_0$ and a parabola when $x_0^2 = 3y_0$ (see figure 6.1). It is easy to obtain the corresponding equation

$$x^2 + ty^2 - (3 + t)y - 1 = 0 \tag{6.5}$$

where t will be determined by the the point (x_0, y_0):

$$t = \frac{3z - x_0^2}{\frac{4}{9} - \frac{5}{3}z + z^2} \tag{6.6}$$

The equation (A.113) can be rewritten as

$$x^2 + t(y - \frac{3+t}{2t})^2 = \frac{9 + 10t + t^2}{4t} \tag{6.7}$$

The above equation corresponds to an ellipse if and only if $t > 0$ ($3z > x_0^2$). Two axes for the ellipse are

$$\hat{a} = \frac{\sqrt{9 + 10t + t^2}}{2t} \qquad \hat{b} = \frac{\sqrt{9 + 10t + t^2}}{2\sqrt{t}} \tag{6.8}$$

So $\zeta = \frac{2t}{\sqrt{1+10t+9t^2}}$ and

$$\frac{d\zeta}{dy_0} = \frac{d\zeta}{dt}\frac{t}{y_0} \approx \frac{2}{3}\frac{27}{4} = \frac{9}{2} \tag{6.9}$$

which justify our conjecture. Also note that (6.8) implies $\hat{b} \approx \hat{a}^{1/2}$ and this shows the minor axes has infinite second moment as well.

The conclusion also holds if the first four points are perturbed by ε within a square of size h^2 around its initial positions.

6.4. Elliptic regression (general case)

To prove the infinite moment for a model containing arbitrary number of points ($n \geqslant 5$), we will modify our constructions as follows: we place $n-4$ points in $\mathbb{B}_3$ and choose h to be extremely small (i.e. $h = 10^{-9}/n^2$) so that the type of best fitting curve is not changed. For every fixed points in $\mathbb{B}_1, \ldots, \mathbb{B}_4$ and the fixed x-coordinate x_n of the last point in $\mathbb{B}_5$, we will examine how the best fitting ellipse changes a parabola as the y-coordinate y_n of the last point changes from $-1/3+h$ to y^*. As the last point further moves down, the best fitting curve that fits to the points changes to a hyperbola for which our geometric parameters are undefined (hyperbola does not have semi axes).

Next we propose to describe ellipse by the following parameters: $p_1 = C_x, p_2 = 1/(a+C_y), p_3 = a - C_y, p_4 = \frac{b^2}{a}$ and $p_5 = \alpha$. When the best fitting conic is a parabola,

A, C_y reach infinity and therefore we define $p_2 = 0$. Let us set $\zeta = p_2$. Also note that $p_3 \approx 1/3$, $p_4 \approx 3/2$(see section A.7 in appendix). It is quite clear that derivative of ζ does not exist when the parameters correspond to a hyperbola because the parameters are undefined. So we have a modified Regularity Lemma below:

LEMMA 6.2. ***[Regularity]*** *For any fixed values (x_i, y_i), $1 \leq i \leq n-1$, and x_n, as above, the function $\zeta(y_n)$ is differentiable and its derivative is bounded when $\zeta > 0$. Furthermore, $\zeta(y_n)$ has bounded one sided derivative when $\zeta = 0$. i.e. $|\zeta'^{+}(y_n)| \leq D$ for some constant $D > 0$. Here D may depend on n and h but not on the fixed coordinates (x_i, y_i).*

It is enough to restrict the region of parameters corresponding to the best fitting curve to

(6.10)

$$\Omega = \left\{ |p_1| \leq 100h, 0 \leq\ p_2 \leq 100h, |p_3 - \frac{1}{3}| \leq 100h, |p_4 - \frac{3}{2}| \leq 100h \quad \text{and} \quad |p_5| \leq 100h \right\}$$

By implicit differentiation of $\mathcal{F}(p_1, p_2, p_3, p_4, p_5)$, we see that all elements in the Hessian Matrix are bounded (see section (A.7) for detail):

(6.11)

$$H = \begin{pmatrix} \frac{232}{325} + \frac{8}{13}n & -\frac{20}{39} + \frac{4}{39}n & \frac{60}{13} - \frac{12}{13}n & \frac{40}{39} - \frac{8}{39}n & -\frac{236}{65} - \frac{12}{13}n \\ -\frac{20}{39} + \frac{4}{39}n & \frac{13162}{2925} + \frac{2}{117}n & -\frac{682}{325} - \frac{2}{13}n & -\frac{6356}{2925} - \frac{4}{117}n & \frac{10}{13} - \frac{2}{13}n \\ \frac{60}{13} - \frac{12}{13}n & -\frac{682}{325} - \frac{2}{13}n & -\frac{232}{325} + \frac{18}{13}n & \frac{116}{325} + \frac{4}{13}n & -\frac{90}{13} + \frac{18}{13}n \\ \frac{40}{39} - \frac{8}{39}n & -\frac{6356}{2925} - \frac{4}{117}n & \frac{116}{325} + \frac{4}{13}n & \frac{2728}{2925} + \frac{8}{117}n & \frac{-20}{13} + \frac{4}{13}n \\ -\frac{236}{65} - \frac{12}{13}n & \frac{10}{13} - \frac{2}{13}n & -\frac{90}{13} + \frac{18}{13}n & -\frac{20}{13} + \frac{4}{13}n & \frac{154}{13} + \frac{18}{13}n \end{pmatrix} + \chi$$

where each element in χ is a small quantity (that can be made as small as we want by further decreasing h). In addition, each leading principal minor of the first

matrix above is positive for $n \geqslant 5$:

$$
\begin{aligned}
M_1 &= \frac{232}{325} + \frac{8}{13}n \\
M_2 &= \frac{934528}{316875} + \frac{21952}{7605}n \\
M_3 &= -\frac{28889344}{316875} + \frac{1277056}{38025}n \\
M_4 &= -\frac{5832704}{950625} + \frac{557056}{316875}n \\
M_5 &= -\frac{8388608}{950625} + \frac{2097152}{950625}n
\end{aligned}
$$

Hence the objective function $\mathcal{F}$ is convex within the compact domain Ω and has unique minimum. Let $\hat{P} = (\hat{p_1}, \hat{p_2}, \hat{p_3}, \hat{p_4}, \hat{p_5})$ be the unique minimum satisfying

$$\mathcal{F}_{p_i}(\hat{p_1}, \hat{p_2}, \hat{p_3}, \hat{p_4}, \hat{p_5}) = 0 \qquad i = 1, 2, 3, 4, 5$$

further differentiate each derivative with respect to y_n gives

$$H \centerdot \gamma + \eta = 0$$

where $\gamma = (\hat{p_1}', \hat{p_2}', \hat{p_3}', \hat{p_4}', \hat{p_5}')^T$ (when $\hat{P}$ corresponds to a parabola, $\hat{p_2}', \hat{p_3}'$ and $\hat{p_4}'$ are replaced by one-sided derivatives) and $\eta = (\mathcal{F}_{p_1 y_n}, \mathcal{F}_{p_2 y_n}, \mathcal{F}_{p_3 y_n}, \mathcal{F}_{p_4 y_n}, \mathcal{F}_{p_5 y_n})^T_{\hat{P}}$. Note that each element of η is uniformly bounded on the compact domain Ω. Therefore, we have $\|\gamma\| \leqslant D$ and $\zeta(y_n)' = \hat{p_2}' < D$. By the regularity lemma,

$$\mathbb{E}(|\hat{a} + \hat{C_y}|) = \infty \tag{6.12}$$

Since $a - C_y \approx 1/3$ and $\frac{b^2}{a} \approx 3/2$, we conclude that$a, b^2, C_x$ and C_y do not have finite moments.

Bibliography

[1] Geometric product specication (gps) acceptance and switzerland (2001) representation test for coordinate measuring machines (cmm) part 6: Estimation of errors in computing gaussian associated features. *Int'l Standard ISO 10360-6. IS, Geneva,*, 2001.

[2] N. Chernov A. Al-Sharadqah and Q. Huang. Errors-in-variables regression and the problem of moments. *Brazilian Journal of Probability and Statistics, to appear*, 2012.

[3] R. J. Adcock. Note on the method of least squares. *Analyst, London*, 4:183–184, 1877.

[4] R. J. Adcock. A problem in least squares. *Analyst, London*, 5:53–54, 1878.

[5] A. Albano. Representation of digitized contours in terms of conic arcs and straight-line segments. *Computer Graphics and Image Processing*, 3:23–33, 1974.

[6] T. W. Anderson. Estimation of linear functional relationships: Approximate distributions and connections with simultaneous equations in econometrics. *J. R. Statist. Soc.B*, 38:1–36, 1976.

[7] R. H. Biggersta. Three variations in dental arch form estimated by a quadratic equation. *J.Dental Res*, 1509, 1972.

[8] F. L. Bookstein. Fitting conic sections to scattered data. *Computer Graphics and Image Processing*, 9:56–71, 1979.

[9] W. H. Breyer. *CRC Standard Mathematical Tables and Formulas*. CRC Press, Boca Raton, FL, 28 edition, 1987.

[10] R. J. Carroll, D. Ruppert, and L. A. Stefansky. *Measurement Error in Nonlinear Models*. Chapman & Hall, London, 1st edition, 1995.

[11] R. J. Carroll, D. Ruppert, L. A. Stefansky, and C. M. Crainiceanu. *Measurement Error in Nonlinear Models: A Modern Perspective*. Chapman & Hall, London, 2nd edition, 2006.

[12] N. Chernov. *Circular and linear regression: Fitting circles and lines by least squares*, volume 117. Chapman & Hall/CRC, 2010.

[13] N. Chernov. *Circular and linear regression: Fitting circles and lines by least squares*, volume 117. Chapman & Hall/CRC, 2010.

[14] N. Chernov. Fitting circles to scattered data: parameter estimates have no moments. *METRIKA*, 73:373–384, 2011.

[15] N. Chernov and C. Lesort. Least squares fitting of circles. *J. Math. Imag. Vision*, 23:239–251, 2005.

[16] N. Chernov and H. Ma. Least squares fitting of quadratic curves and surfaces. *Computer Vision, Editor S. R. Yoshida, Nova Science Publishers*, pages 285–302, 2011.

[17] W. R. Cook. On curve fitting by means of least squares. *Philos. Mag. Ser. 7*, 12:1025–1039, 1931.

[18] W. E. Deming. The application of least squares. *Philos. Mag. Ser. 7*, 11:146–158, 1931.

[19] http://www.mathworld.wolfram.com/CantorDiagonalMethod.html.

[20] M. G. Kendal. Regression, structure, and functional relationships, part i. *Biometrica.*, 38:11–25, 2051.

[21] M. G. Kendal. Regression, structure, and functional relationships, part ii. *Biometrica.*, 39:96–108, 2052.

[22] C. H. Kummell. Reduction of observation equations which contain more than one observed quantity. *Analyst, London*, 6:97–105, 1879.

[23] K Levenberg. A method for the solution of certain problems in least squares. *Quart. Appl. Math*, 2:164–168, 1944.

[24] E. Malinvaud. *Statistical Methods of Econometrics*. North Holland Publ. Co, Amsterdam, 3rd edition, 1980.

[25] D Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math*, 11:431–443, 1963.

[26] Y. Nievergelt. Hyperspheres and hyperplanes fitted seamlessly by algebraic constrained total least-squares. *Linear Algebra Appl.*, 331:43–59, 2001.

[27] Y. Nievergelt. A finite algorithm to fit geometrically all midrange lines, circles, planes, spheres, hyperplanes, and hyperspheres. *J. Numerische Math.*, 91:257–303, 2002.

[28] Y. Nievergelt. Fitting conics of specific types to data. *Linear Algebra Appl.*, 378:1–30, 2004.

[29] S.C. Pei and J.H. Horng. Optimum approximation of digital planar curves using circular arcs. *Pattern Recogn*, pages 383–388, 1996.

[30] N. Chernov Q. Huang and H. Ma. Does the best fitting curve always exist? *ISRN Probability and Statistics*, 2011.

[31] E. Zelniker and V. Clarkson. A statistical analysis of the Delogne-Kåsa method for fitting circles. *Digital Signal Proc.*, 16:498–522, 2006.

APPENDIX A

APPENDIX

THEOREM A.1. *Given a closed set S_1 and a compact S_2 on the plane, the distance between S_1 and S_2 is defined by*

$$\text{dist}(S_1, S_2) = \inf_{P_1 \in S_1, P_2 \in S_2} \text{dist}(P_1, P_2) \tag{A.1}$$

If one set (say, S_1) is closed and the other (S_2) is compact, the infimum in (2.4) can always be replaced by a minimum.

PROOF. It is possible to find two sequences $P_{1i} \in S_1$, $P_{2i} \in S_2$ such that

$$\text{dist}(P_1^{(i)}, P_2^{(i)}) \to \inf_{P_1 \in S_1, P_2 \in S_2} \text{dist}(P_1, P_2)$$

By the compactness of S_2, let us assume that $P_2^{(i)} \to P_2' \in S_2$. The sequence $P_1^{(i)}$ has a subsequence either diverging to infinity or converging to a finite limit within S_1 denoted by P_1'. However, if the subsequence $P_1^{(i)}$ moves to infinity, $\text{dist}(P_1^{(i)}, P_2^{(i)}) \to \infty$. For convenience, let us assume $P_1^{(i)} \to P_1'$. Now we see that

$$\text{dist}_{P_1^{(i)} \in S_1, P_2^{(i)} \in S_2}(P_1^{(i)}, P_2^{(i)}) \to \text{dist}_{P_1' \in S_1, P_2' \in S_2}(P_1^{(i)}, P_2^{(i)}) = \inf_{P_1 \in S_1, P_2 \in S_2} \text{dist}(P_1, P_2)$$

which implies

$$\inf_{P_1 \in S_1, P_2 \in S_2} \text{dist}(P_1, P_2) = \min_{P_1 \in S_1, P_2 \in S_2} \text{dist}(P_1, P_2)$$

. □

THEOREM A.2. *Given a sequence of sets S_n and a set S, we have*

$$\text{dist}_{\text{B}}(S_n, S) = \sum_{k=1}^{\infty} 2^{-k} \text{dist}_{\text{H}}(S_n, S; R_k) \to 0 \qquad \text{as} \qquad n \to \infty \tag{A.2}$$

if and only if

$$\text{dist}_{\text{H}}(S_n, S; R_k) = \max\Big\{ \sup_{P \in S_n \cap R} \text{dist}(P, S), \sup_{Q \in S \cap R} \text{dist}(Q, S_n) \Big\} \to 0 \tag{A.3}$$

for each R_k *(*$R_k = \{-k \le x \le k, \quad -k \le y \le k\}$*).*

PROOF. Let us first suppose the $\mathrm{dist}_{\mathrm{B}}(S_n, S)$ has a limit 0. If

$$\mathrm{dist}_{\mathrm{H}}(S_n, S; R_j) \not\to 0 \quad \text{as} \quad n \to \infty$$

for some j, there exists a subsequence

$$\mathrm{dist}_{\mathrm{H}}(S_{n_i}, S; R_j) \to c > 0 \quad \text{or} \quad +\infty \quad \text{as} \quad n_i \to \infty$$

Then

$$\mathrm{dist}_{\mathrm{B}}(S_{n_i}, S) = \sum_{k=1}^{\infty} 2^{-k} \mathrm{dist}_{\mathrm{H}}(S_{n_i}, S; R_k) > 2^{-j} c$$

for any n_i or

$$\mathrm{dist}_{\mathrm{B}}(S_{n_i}, S) \to +\infty$$

Which contradicts our assumptions.

Next, let us suppose

$$\text{(A.4)} \qquad \mathrm{dist}_{\mathrm{H}}(S_n, S; R_k) = \max\Big\{ \sup_{P \in S_n \cap R_k} \mathrm{dist}(P, S), \sup_{Q \in S \cap R_k} \mathrm{dist}(Q, S_n) \Big\} \to 0,$$

for any R_k. Note that

(A.5)

$$\begin{aligned} \sup_{P \in S_n \cap R_{k+1}} \mathrm{dist}(P, S) &= \mathrm{dist}_{P \in S_n \cap R_{k+1}, Q \in S}(P, Q) \\ &< \mathrm{dist}_{P \in S_n \cap R_{k+1}, P' \in S_n \cap R_k}(P, P') + \mathrm{dist}_{Q \in S, P' \in S_n \cap R_k}(P', Q) \\ &< \sqrt{2}(2k+1) + \sup_{P \in S_n \cap R_k} \mathrm{dist}(P, S) \end{aligned}$$

(The farthest possible distance between a point $Q_1 \in R_k$ and another $Q_2 \in R_{k+1}$ is $\sqrt{2}(2k+1)$, See Figure A.1). It is easy to verify that $\sum_{k=j}^{\infty} 2^{-k+1/2}(k-j)(2k+1)$ has a finite limit for any fixed j. So for any $\varepsilon > 0$, there exists a $p > j$ such that $\sum_{k=p}^{\infty} 2^{-k+1/2}(k-j)(2k+1) < \varepsilon$. Since $\mathrm{dist}_{\mathrm{H}}(S_n, S; R_p) \to 0$, $\mathrm{dist}_{\mathrm{H}}(S_n, S; R_p) < \varepsilon$

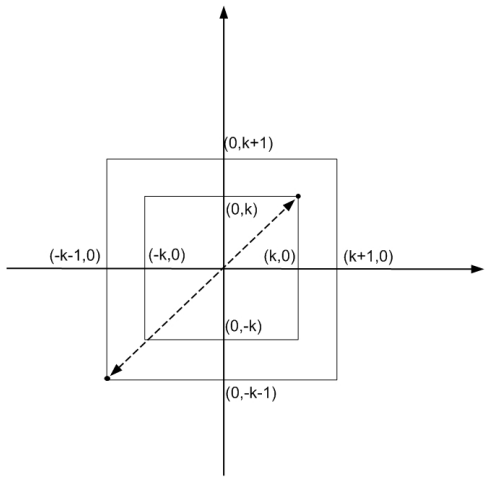


Figure A.1: The farthest possible distance between one point in the inner square and the other in the outer square.

for sufficiently large n. Then

$$\begin{aligned}
(\text{A.6}) \quad \operatorname{dist}_{\mathrm{B}}(S_n, S) &= \sum_{k=1}^{p-1} 2^{-k} \operatorname{dist}_{\mathrm{H}}(S_n, S; R_k) + \sum_{k=p}^{\infty} 2^{-k} \operatorname{dist}_{\mathrm{H}}(S_n, S; R_k) \\
&< \sum_{k=1}^{p-1} 2^{-k}\varepsilon + \sum_{k=p}^{\infty} 2^{-k}(\operatorname{dist}_{\mathrm{H}}(S_n, S; R_p) + (k-p)\sqrt{2}(2k+1)) \\
&< \sum_{k=1}^{p-1} 2^{-k}\varepsilon + \sum_{k=p}^{\infty} 2^{-k}(\operatorname{dist}_{\mathrm{H}}(S_n, S; R_p) + (k-j)\sqrt{2}(2k+1)) \\
&< \varepsilon + \varepsilon + \varepsilon \\
&< 3\varepsilon
\end{aligned}$$

indicating that $\operatorname{dist}_{\mathrm{B}}(S_n, S) \to 0$. The proof of our claim is complete. □

A.1. No local minima for five distinct points

Suppose five points $P_1, \ldots, P_5$ are given. Then there exists an interpolating object S_{best} (quadratic curve or line). Thus the objective function $\mathcal{F}$ takes it global minimum $\mathcal{F}(S_{\text{best}}) = 0$. A local minimum of $\mathcal{F}$ would be a conic S such that $\mathcal{F}(S) > 0$ (i.e., S

does not interpolate our points), but $\mathcal{F}(S) \leq \mathcal{F}(S')$ for any conic S' sufficiently close to S (in the sense of our section (2.2)). It turns out that local minima does not exist.

LEMMA A.1. *For five distinct points that can be interpolated by an ellipse, if one of them changes its position by a sufficiently small amount, there exists another ellipse interpolating the new set of five points.*

PROOF. Suppose the five points are located at (x_i, y_i) $i = 1, ..., 5$, which uniquely determine an ellipse:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + 1 = 0 \tag{A.7}$$

where $B^2 - 4AC > 0$ (see chapter). Then

$$Ax_i^2 + Bx_iy_i + Cy_i^2 + Dx_i + Ey_i + 1 = 0 \quad i = 1, \ldots, 5$$

Suppose

$$Q = -\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}^T$$

and

$$X = \begin{pmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 \\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 \\ x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 \\ x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 \\ x_5^2 & x_5y_5 & y_5^2 & x_5 & y_5 \end{pmatrix}$$

The parameter vector $\theta = (A, B, C, D, E)^T$ is uniquely determined by $\theta = X^{-1}Q$.
If we consider the first four points (x_i, y_i) $i = 1, 2, 3, 4$ as fixed and only move the fifth (x_5, y_5) point within a region $S = \{(x_5, y_5) | \det(X) \neq 0\}$. The parameters A, B and C can be considered as functions of (x_5, y_5)(defined on S, denoted by f_1, f_2 and f_3 respectively. Also note that $4f_1(x, y)f_3(x, y) - f_2^2(x, y) > 0$ when the equation corresponds to an ellipse. Let $F(x, y) = 4f_1(x, y)f_3(x, y) - f_2^2(x, y)$. Apparently, $F(x, y)$ is a continuous function on S. Suppose $I = (F(x_5, y_5) - \delta, F(x_5, y_5) + \delta)$ where $\delta < F(x_5, y_5)$. Then $(x_5, y_5) \in F^{-1}(I)$ which is an open set in R. Furthermore,

(x_5, y_5) has an open neighborhood U contained in $F^{-1}(I)$ and $F(U) \subset I$. Therefore for any perturbation of (x_5, y_5) within U, the conic interpolating all points is still an ellipse.

□

LEMMA A.2. *Suppose an ellipse* $\mathbb{E}$ *passes through* 4 *distinct points* P_i, $1 \leq i \leq 4$, *and* T *denotes the tangent line to the ellipse at* P_1. *Then* $\mathbb{E}$ *is uniquely determined.*

PROOF. First, choose an appropriate coordinate system so that the given tangent line passes through the origin and the P_1 is in the upper half of the coordinate system but not at the origin. Also assume the tangent line has a slope $0 < k < \infty$. Let x-coordinate of Q_1 equals to $1/2$ which indicates $y_1 = k/2$. Then the orthogonal vector $(2Ax_1 + By_1 + D, 2Cy_1 + Bx_1 + E)$ at (x_1, y_1) satisfies

$$2Ax_1 + By_1 + D = -k(2Cy_1 + Bx_1 + E) \quad or \quad A + Bk + Ck^2 + D + kE = 0 \tag{A.8}$$

Consider the system of equations for the parameter vector of $\mathbb{E}$ below.

$$X\theta = \begin{pmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 \\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 \\ x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 \\ x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 \\ 1 & k & k^2 & 1 & k \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \\ E \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \tag{A.9}$$

The matrix on the left is invertible. Here is the reason: There is a unique conic passing through five distinct pointsP_i, $1 \leq i \leq 4$ and $(1, k)$. So the system of equations of θ

$$X\theta = -(1, 1, 1, 1, 1)^t \tag{A.10}$$

has an unique solution which implies $det(X) \neq 0$. Therefore there is an unique ellipse $\mathbb{E}$, which is determined by (A.9). □

THEOREM A.3. *In a model of fitting any conics (ellipses, parabola and hyperbola) to* 5 *distinct points on* $\mathbb{R}^2$, *if all five points are interpolated by a hyperbola, there*

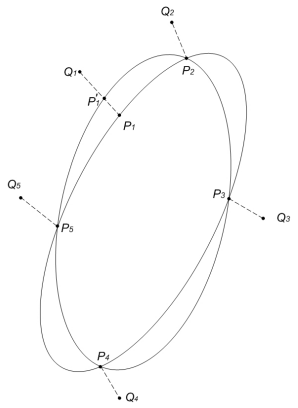


Figure A.2: Five different projections

doesn't exist a local minimum in the space of ellipses. More precisely, for any given ellipse fitted to the points, one can find another ellipse that produces a smaller sum of squares of orthogonal distances.

PROOF. Let $Q_1, \ldots, Q_5$ be five points on a hyperbola and $\mathbb{E}$ a given ellipse. It's necessary to discuss two possible situations:

(I) The points P_1 through P_5 have different projections on $\mathbb{E}$ (see Figure A.2.)

Let's denoted by $P_1, \ldots, P_5$ the projections of $Q_1, \ldots, Q_5$. By lemma (A.1), there is an open neighborhood of P_1 denoted by $\mathbb{U}$ in which any perturbation of P_1 will still lead to an ellipse. Let us pick a arbitrary point $P_1^{'} \in \mathbb{U}$ between P_1 and Q_1. Then one can fit $P_1^{'}, P_2, \ldots, P_5$ by another ellipse for which

$$[\text{dist}(Q_1, P_1^{'})]^2 + \sum_{i=2}^{5} \text{dist}(Q_i, P_i)]^2 < \sum_{i=1}^{5} \text{dist}(Q_i, P_i)]^2$$

Therefore the new ellipse interpolating the new set of points fits the given points $Q_1, \ldots, Q_5$ better than $\mathbb{E}$.

(II) Two points Q_1 and Q_2 share a projection P_{12} on $\mathbb{E}$.

The rest of projections are three distinct points P_3, P_4 and P_5. It is easy to see that P_1 and P_2 can not be on $\mathbb{E}$ at the same time. Let us suppose P_1 does not belong to $\mathbb{E}$. Pick a point P_6 on $\mathbb{E}$ different from P_{12}, P_3, P_4 and P_5. Next move P_6 to $P_6^{'}$ within

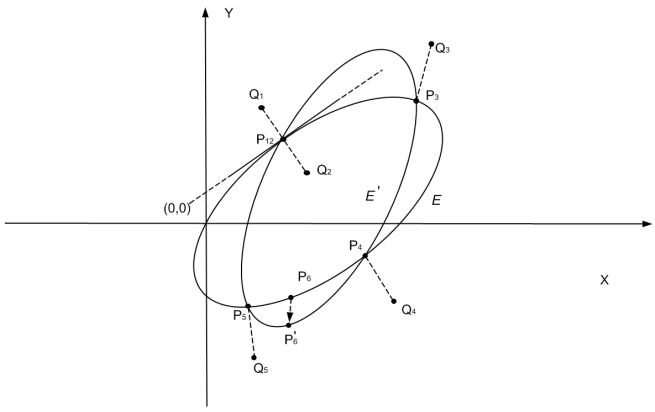


Figure A.3: Two points share an identical projection

a certain neighborhood where the conic interpolating P_{12}, P_3, P_4, P_5 and $P_6^{'}$ is still an ellipse $\mathbb{E}^{'} \neq \mathbb{E}$ (by lemma (A.1)). By lemma (A.2), the new ellipse has different tangent lines at P_{12}, P_3, P_4 and P_5. Then $\mathrm{d}ist(\mathbb{E}^{'}, P_i) < \mathrm{d}ist(\mathbb{E}, P_i)$ $(i = 12, 3, 4, 5)$. So $\mathbb{E}^{'}$ has a smaller sum of squares of distances to given points than $\mathbb{E}$ (see figure A.3).

(III)Q_1, Q_2 share a projection P_{12}; Q_3, Q_4 share a projection P_{34}

Suppose P_5 has a projection Q_5. Pick two points Q_6 and Q_7 on $\mathbb{E}$ other than Q_{12}, Q_{34} and Q_5. Same argument will apply if one moves Q_5 toward P_5 within its small neighborhood.

Remember we assume all five original points are on a hyperbola. If three of them have the same projection, the best fitting conic can not interpolate three distinct points on a straight line. So by checking the above situation, the proof is complete. □

A.2. Existence of the best fit: specific models

Here we provide sketchy analytic proofs as alternative to topological proofs given earlier.

THEOREM A.4. *Let B be a given compact set (Euclidean space) containing all the data points. Then the 'enlarged' space Ω of ellipses, parabolas, lines(including rays*

and line segments, singletons), and pairs of parallel lines intersecting B is compact with respect to the topology defined on the Ω (see section (2.3)).

PROOF. Let S_i be a sequence of curve objects (ellipses, parabolas, lines, pairs of parallel lines and singletons) that intersect B. Denote by $d_{\max}$ the longest distance from the origin to a point in the box B. Then each line in the space Ω is uniquely determined by the direction of the normal vector to the line $0 \leqslant \omega < 2\pi$ and the distance from the origin to the line $0 \leqslant d \leqslant d_{\max}$. So every sequence of lines in the space contains a convergent subsequencein Ω.

If there are infinitely many ellipses, each of them is represented by its focuses f_i, g_i and the semi minor axis b_i. Then there is a subsequence of ellipses satisfying one of the following conditions:

(i)$f_i \to f_0$ and $g_i \to \infty$;

Without loss of generality, let's assume a sequence of ellipses with one focus fixed on the origin and another one denoted by a polar coordinate $(\theta_i, 2k_i)$ moving to the infinity. Since $0 \leqslant \theta_i < 2\pi$, let's just assume $\theta_i \to \theta_0$. Such ellipses with length of semi-minor axis b_i can be represented by a family of equations:

$$(k_i^2 \sin^2\theta_i + b_i^2)x^2 + (k_i^2 \cos^2\theta_i + b_i^2)y^2 - xyk_i^2 \sin 2\theta_i - 2k_i b_i^2(x\cos\theta_i + y\sin\theta_i) - b_i^4 = 0 \tag{A.11}$$

(1)$b_i^2/k_i \to c \in [0, \infty)$

After dividing both side of (1) by k_i^2 and set $k_i \to \infty$ which is already implied by the assumption $g_i \to \infty$, (1) becomes

$$x^2 \sin^2\theta_0 + y^2 \cos^2\theta_0 - 2xy \sin\theta_0 - 2c(x\cos\theta_0 + y\sin\theta_0) - c^2 = 0 \tag{A.12}$$

Since $(-2\sin\theta_0)^2 - 4\sin^2\theta_0\cos^2\theta_0 = 0$, the above equation corresponds to a parabola when the above equation corresponds to a parabola if $c \neq 0$ or a ray starting at the

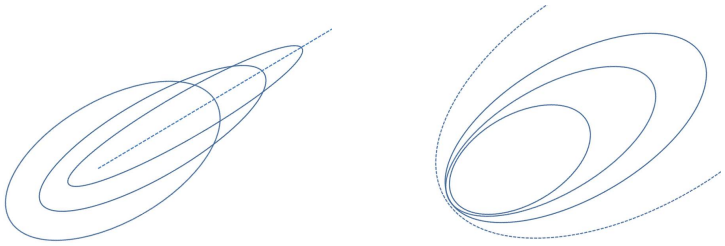


Figure A.4: ellipses converging to a ray or a parabola

limiting vertex:

$$\begin{aligned}&\lim((\sqrt{b_i^2+k_i^2}-k_i)\cos\theta_i,(\sqrt{b_i^2+k_i^2}-k_i)\sin\theta_i)\\&=((\sqrt{1+c}-1)\cos\theta_0,\sqrt{1+c}-1)\sin\theta_0)\\&=(0,0)\end{aligned}$$

Thus we get the equation for the ray

$$y=x\tan\theta_0\begin{cases}x\in[0,\infty), & \text{if } \theta_0\in[0,\frac{\pi}{2}]\bigcup[\frac{3\pi}{2},2\pi];\\ x\in(\infty,0), & \text{if } \theta_0\in(\frac{\pi}{2},\frac{3\pi}{2});\end{cases}$$

or

$$x=0\begin{cases}y\in[0,\infty), & \text{if } \theta_0=\frac{pi}{2};\\ y\in(\infty,0], & \text{if } \theta_0=\frac{3pi}{2};\end{cases}$$

(2)$b_i^2/k_i\to\infty$

Apparently $b_i\to\infty$ and $k_i/b_i^2\to 0$. Dividing both side of (1) by b_i^4 yields

$$\begin{aligned}(\text{A.13})\quad&((k_i^2\sin^2\theta_i+b_i^2)x^2+(k_i^2\cos^2\theta_i+b_i^2)y^2\\&-xyk_i^2\sin 2\theta_i-2k_ib_i^2(x\cos\theta_i+y\sin\theta_i))/b_i^4-1=0\end{aligned}$$

The first fraction on the left side of (3) converges to 0 as $i\to\infty$ and the equation turns into $-1=0$ which does not represents any curve intersecting with B. Therefore, subsequence of ellipses satisfying condition (i) could only converge to a parabola

or a ray.

(ii)$f_i, g_i \to \infty$

Let's denote by a_i and b_i the length of two semi axes of an ellipse.

(1)$b_i \to b_0 < \infty$

If $dist(f_i, g_i) \to d$ (i.e. $2\sqrt{a_i^2 - b_i^2} \to d$), all ellipses have bounded semi axes. Since they intersect with B, it can not happen that both focuses $f_i, g_i \to \infty$. So $dist(f_i, g_i) \to \infty$,let's consider the canonic equation of an ellipse

$$(X - X_c)^t A(X - X_c) = 1 \tag{A.14}$$

where A is a symmetric matrix. The Singular value decomposition of A:

$$U^t \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} U = \lambda_1 u^t u + \lambda_2 v^t v \tag{A.15}$$

Note that u and v are the unit eigenvectors corresponding to the λ_1 and λ_1. Substituting (5) into (4) yields

$$(x - x_c, y - y_c)(\lambda_1 u^t u + \lambda_2 v^t v) \begin{pmatrix} x - x_c \\ y - y_c \end{pmatrix} \tag{A.16}$$

Suppose $u = (u_1, u_2)$ and $v = (v_1, v_2)$. After expansion, (6) turns into

$$\begin{aligned} &\lambda_1(u_1 x + u_2 y)^2 + \lambda_2(v_1 x + v_2 y)^2 \\ &\quad - 2[u_1\lambda_1(u_1 x_c + u_2 y_c) + v_1\lambda_2(v_1 x_c + v_2 y_c)]x \\ &\quad - 2[u_2\lambda_1(u_1 x_c + u_2 y_c) + v_2\lambda_2(v_1 x_c + v_2 y_c)]y \\ &\quad + \lambda_1(u_1 x_c + u_2 y_c)^2 + \lambda_2(v_1 x_c + v_2 y_c)^2 = 1 \end{aligned} \tag{A.17}$$

Since the length of the major axis $2a_i = \frac{2}{\sqrt{\lambda_1}}$ grows to infinity,$\lambda_1 \to 0$. Next, let's recognize the following facts about the subsequence.

i.The distance from the major axis to the origin will not grow to infinity because the length of the minor axis is bounded and the subsequence can not leave the box B under the assumption.

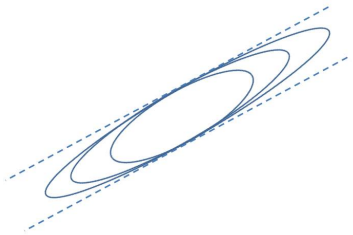


Figure A.5: Stretch ellipses to a pair of parallel lines

ii.Suppose $u_1x + u_2y + d_1 = 0$ and $v_1x + v_2y + d_2 = 0$ pass through the center of the ellipse (x_c, y_c). Since $\|u_i\| = 1$, d_1 and d_2 are the distances from the minor and major axes to the origin.

iii. Because of i, $d_2 \to d_2^{'} < \infty$. Furthermore, if we draw a line through the origin parallel to the minor axis, the length of the section between the line and the minor axis is d_1 which must be shorter than the major axis.

By above facts, we have

$$\lim_{\lambda_1 \to 0} \frac{d_1}{2a} = \lim_{\lambda_1 \to 0} \frac{-(u_1x_c + u_2y_c)}{2/\sqrt{\lambda_1}} = t \in [0, \frac{1}{2}]$$

Then

$$\lim_{\lambda_1 \to 0} \lambda_1(u_1x_c + u_2y_c) = 0$$

$$\lim_{\lambda_1 \to 0} \lambda_1(u_1x_c + u_2y_c)^2 = 4t^2 = c \in [0, 1]$$

Suppose $u \to u^{'}$ and $v \to v^{'}$ Therefore, the equation (7) will converge to

$$(u_1^{'}x + u_2^{'}y)^2 + 2v_1^{'}d_2^{'}x + 2v_2^{'}d_2^{'}y + d_2^{'2} = \frac{1-c}{4}b_0^2$$

or equivalently

$$(u_1^{'}x + u_2^{'}y + d_2^{'})^2 = \frac{1-c}{4}b_0^2$$

The equation represents a pair of parallel lines $u_1^{'}x + u_2^{'}y + d_2^{'} \pm \frac{\sqrt{1-c}}{2}b_0 = 0$ or a line $u_1^{'}x + u_2^{'}y + d_2^{'} = 0$ if $c = 1$ or $b_0 = 0$. See Figure 2.

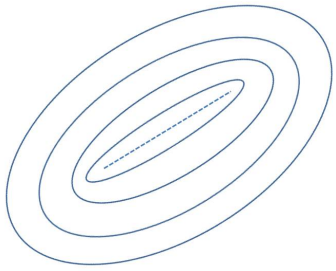


Figure A.6: Shrink ellipses to a line segment

(2)$b_i \to \infty$ As both axes stretch to infinity, the elliptic segment inside the box would be straighten up and the subsequence of ellipses converges to a line.

(iii)$f_i \to f_0 \quad g_i \to g_0$

Any sequence of ellipses satisfying (iii) in the space Ω must have bounded length of axes. The subsequence of ellipses could either converge to an ellipse or a singleton ((both axes shrink to a point). However, if $2a_i \to dist(f_0, g_0) > 0$ (equivalent to the length of the minor axis $2b = \frac{2}{\sqrt{\lambda_2}} \to 0$, the limit of the subsequence is a line segment.See figure 3. Suppose the $f_0 = (f_x, f_y)$ and $g_0 = (g_x, g_y)$. Then the equation (7) converges to

$$u_1^{'}x + u_2^{'}y + d_2^{'} = 0 \quad x \in [\min(f_x, g_x), \max(f_x, g_x)]$$

if $u_1^{'} \neq 0$ or otherwise

$$u_2^{'}y + d_2^{'} = 0 \quad y \in [\min(f_y, g_y), \max(f_y, g_y)]$$

Next, let's prove that every sequence of parabolas or parallel lines contains a convergent subsequence. If there are infinitely many parabolas in the sequence, each of them uniquely determined by its directrix line and focus point. Given a sequence of parabolas with the focus $f_i = (k_i \cos\theta_i, k_i \sin\theta_i)$ and the directrix line $x\cos\alpha_i + y\sin\alpha_i - d_i = 0$ where α_i is the direction of its normal vector and d_i is its distance

to the origin, the parabola is naturally described by the equation

$$(A.18)\quad x^2\sin^2\alpha_i + y^2\cos^2\alpha_i - xy\sin 2\alpha_i - 2(k_i\cos\theta_i - d_i\cos\alpha_i)x - 2(k_i\sin\theta_i - d_i\sin\alpha_i)y + k_i^2 - d_i^2 = 0$$

Suppose $\alpha_i \to \alpha_0$, $\theta_i \to \theta_0$

(i)$d_i \to d_0 \quad k_i \to k_0$

If f_i goes onto the directrix line, the points with equal distances to the directrix and the focus form a ray orthogonal to the directrix line and starting from $(k_0\cos\theta_0, k_0\sin\theta_0)$ (e.x.$d_i, k_i \to 0$). Otherwise, the sequence converges to a parabola.

(ii)$d_i \to \infty$, $k_i \to k_0$ or $d_i \to d_0$, $k_i \to \infty$

(8) is equivalent to

$$(A.19)\quad [x^2\sin^2\alpha_i + y^2\cos^2\alpha_i - xy\sin 2\alpha_i - 2(k_i\cos\theta_i - d_i\cos\alpha_i)x - 2(k_i\sin\theta_i - d_i\sin\alpha_i)y + k_i^2]/d_i^2 - 1 = 0$$

which turns into $-1 = 0$ as $d_i \to \infty$. Similar argument applies to $k_i \to \infty$while $d_i \to d_0$. So such sequence does not exist in the space of parabolas intersecting B.

(iii)$d_i \to \infty$, $k_i \to \infty$

(1)$k_i - d_i \to \infty$

If we divide both side of (8) by $(k_i + d_i)(k_i - d_i)$, we have following equation:

$$(A.20)\quad \frac{x^2\sin^2\alpha_i + y^2\cos^2\alpha_i - xy\sin 2\alpha_i}{(k_i+d_i)(k_i-d_i)} - \frac{2(k_i\cos\theta_i - d_i\cos\alpha_i)x}{(k_i+d_i)(k_i-d_i)} \frac{-2(k_i\sin\theta_i - d_i\sin\alpha_i)y}{(k_i+d_i)(k_i-d_i)} + 1 = 0$$

The first three fractions converge to 0 and it becomes $1 = 0$ which again indicates the absence of such type of sequence.

(2)$k_i - d_i \to Q < \infty$

Since $(k_i - d_i)/d_i \to 0$, $k_i/d_i \to 1$. Then by dividing the second and the third fraction

of above equation and set $d_i \to \infty$,

$$\frac{x^2 \sin^2 \alpha_i + y^2 \cos^2 \alpha_i - xy \sin 2\alpha_i}{k_i + d_i} - \frac{2(k_i/d_i \cos\theta_i - \cos\alpha_i)x}{k_i/d_i + 1} \\ \frac{-2(k_i/d_i \sin\theta_i - \sin\alpha_i)y}{k_i/d_i + 1} + k_i - d_i = 0 \tag{A.21}$$

one can obtain

$$(\cos\theta_0 - \cos\alpha_0)x + (\sin\theta_0 - \sin\alpha_0)y - Q = 0$$

It represents a line when $\theta_0 \neq \alpha_0$.

If $\theta_0 = \alpha_0$, we only need to consider the case $Q = 0$. Without loss of generality, we assume $\theta_i = \alpha_i = 0$ and $k_i = d_i$ for all $i = 1, ..., \infty$. The ellipses are represented by

$$y^2 = 4k_i x \quad i = 1, ..., \infty$$

which reaches the limit $x = 0$ as $k_i \to \infty$. Thus, the sequence of ellipses converges to a line through the origin.

In the end, one can easily investigate a sequence of two parallel lines, rays or line segments in a similar manner and find their limits contained by the space Ω. So the lemma is proved. □

LEMMA A.3. *Let $A_0x + B_0y + C_0 = 0$ and $A_1x + B_1y + C_1 = 0$ ($A_0B_1 \neq A_1B_0$) be asymptotes of some hyperbola. Then the hyperbola has an equation $(A_0x + B_0y + C_0)(A_1x + B_1y + C_1) = a$ ($a \neq 0$).*

PROOF. Geometrically, any hyperbola with its center at (x', y') and angle θ between major axis and horizontal axis can be achieved by transforming a hyperbola in the standard position

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

with shift and rotation. Let us consider the equation in matrix notation:

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \frac{1}{a} \\ \\ \frac{1}{b} \end{pmatrix} \begin{pmatrix} \frac{1}{a} & -\frac{1}{b} \end{pmatrix} \begin{pmatrix} x \\ \\ y \end{pmatrix} = 1 \tag{A.22}$$

Note that the corresponding asymptotes are $x/a \pm y/b = 0$. Applying transformation to the hyperbola equivalently changes (A.22) to

$$\begin{pmatrix} x-x' & y-y' \end{pmatrix} R' \begin{pmatrix} \frac{1}{a} \\ \frac{1}{b} \end{pmatrix} \begin{pmatrix} \frac{1}{a} & -\frac{1}{b} \end{pmatrix} R \begin{pmatrix} x-x' \\ y-y' \end{pmatrix} = 1 \tag{A.23}$$

where $R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Meantime, two asymptotes are geometrically transformed in the same manner. So they have equations

$$\begin{pmatrix} x-x' & y-y' \end{pmatrix} R' \begin{pmatrix} \frac{1}{a} \\ \frac{1}{b} \end{pmatrix} = 0 \tag{A.24}$$

and

$$\begin{pmatrix} \frac{1}{a} & -\frac{1}{b} \end{pmatrix} R \begin{pmatrix} x-x' \\ y-y' \end{pmatrix} = 0 \tag{A.25}$$

which are two factors on the left side of (A.23). So the lemma is proved. □

Next, we will use this lemma in the formal proof of the following theorem.

THEOREM A.5. *Let B be a given bounding box containing all the data points. The 'enlarged' set Ω of ellipses, parabolas, hyperbolas, lines (including rays and line segments, singletons and opposite rays), pairs of parallel lines, pairs of intersecting lines crossing B is compact.*

PROOF. To justify the compactness of a set, by the conclusion in Existence of the best fit, we only need to check if it contains all of its limit points. Again, convergence of sequence of objects to a limit object is defined in . To search for the limit points (objects) analytically, we will instead work with convergence of parameter vectors which implies the convergence of sequence of objects (See section 4.6). Let us begin by looking at a sequence containing infinitely many hyperbolas. By lemma A.3, each hyperbola can be represented by equation:

$$(x\cos\theta_i + y\sin\theta_i + d_i)(x\cos\varphi_i + y\sin\varphi_i + d_i^{'}) = a_i \quad i = 1, 2, \ldots \tag{A.26}$$

with associated asymptotes $x\cos\theta_i + y\sin\theta_i + d_i = 0$ and $x\cos\varphi_i + y\sin\varphi_i + d_i^{'} = 0$ (Lemma (A.3)). By requiring $0 \leqslant \theta_i \leqslant 2\pi$ and $d_i \geqslant 0$, we gain some additional benefits: θ_i, φ_i will be the direction of normal vectors to two asymptotes and $d_i, d_i^{'}$ the distance between the origin and the asymptotes.

The equation (A.26) is equivalent to a quadratic equation in the variable y

$$\begin{aligned}(A.27)\quad & y^2\sin\theta_i\sin\varphi_i + y(x\sin(\theta_i+\varphi_i) + d_i^{'}\sin\theta_i + d_i\sin\varphi_i) \\ & + x^2\cos\theta_i\cos\varphi_i + (d_i^{'}\cos\theta_i + d_i\cos\varphi_i)x + d_i d_i^{'} - a_i = 0\end{aligned}$$

when $\sin\theta_i\sin\varphi_i \neq 0$. The solution exists if and only if the discriminant

$$(A.28)\qquad \Delta = (x\sin(\theta_i-\varphi_i) - d_i^{'}\sin\theta_i + d_i\sin\varphi_i)^2 + 4a_i\sin\theta_i\sin\varphi_i \geqslant 0$$

Since any bounded sequence always has a convergent subsequence, let us just assume $\sin\theta_i\sin\varphi_i \to \sin\theta\sin\varphi$. Note that $\sin\theta_i\sin\varphi_i = 0$ can be avoid because one has freedom of choosing appropriate coordinate system such that $\sin\theta\sin\varphi \neq 0$. So we can stick with the fact $\sin\theta\sin\varphi \neq 0$ and $\sin\theta_i\sin\varphi_i \neq 0$ for all i's in the following proof. Since two asymptotes of hyperbolas can not be parallel, $\theta_i - \varphi_i \neq 0$, π and 2π and we can rewrite (A.28) as

$$(A.29)\qquad (x - \frac{d_i^{'}\sin\theta_i - d_i\sin\varphi_i}{\sin(\theta_i-\varphi_i)})^2 + \frac{4a_i\sin\theta_i\sin\varphi_i}{\sin^2(\theta_i-\varphi_i)} \geqslant 0$$

Let

$$(A.30)\qquad T_i = \frac{d_i^{'}\sin\theta_i - d_i\sin\varphi_i}{\sin(\theta_i-\varphi_i)} \quad \text{and} \quad Q_i = \frac{4a_i\sin\theta_i\sin\varphi_i}{\sin^2(\theta_i-\varphi_i)}$$

We will consider different cases based on the limits of the parameters.

(I) $a_i \to 0$.

(i) If $d_i \to d < \infty$ and $d_i^{'} \to \infty$, (A.26) is equivalent to

$$(A.31)\qquad (x\cos\theta_i + y\sin\theta_i + d_i)(\frac{x\cos\varphi_i + y\sin\varphi_i}{d_i^{'}} + 1) = \frac{a_i}{d_i^{'}}$$

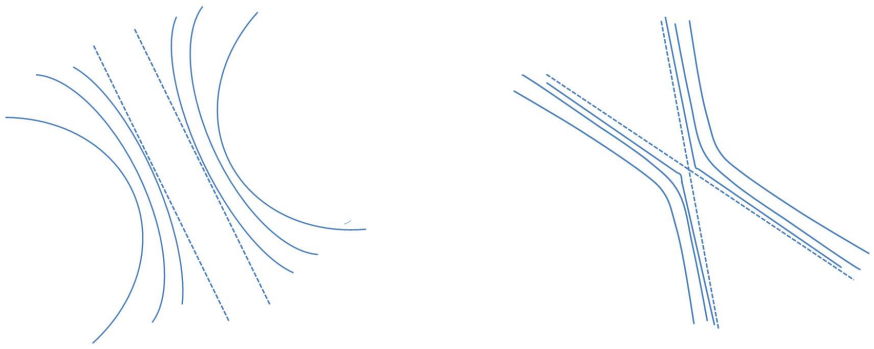


Figure A.7: Hyperbolas converge to a pair of parallel lines or intersecting lines

which approximates to

$$x\cos\theta + y\sin\theta + d = 0 \tag{A.32}$$

after limits are taken on both sides. So the limit object represented by above equation is a straight line.

(ii) If $d_i \to \infty$ and $d_i^{'} \to \infty$,

$$(\frac{x\cos\theta_i + y\sin\theta_i}{d_i} + 1)(\frac{x\cos\varphi_i + y\sin\varphi_i}{d_i^{'}} + 1) = \frac{a_i}{d_i d_i^{'}} \tag{A.33}$$

where the left side converges to 1 while the right side 0 as $i \to \infty$. In fact, the equations corresponds to a sequence of hyperbola moving away from Ω to infinity, which contradict our primary assumption all model objects intersect with Ω.

(iii) If $d_i \to d < \infty$ and $d_i^{'} \to d^{'} < \infty$, (A.26) transforms into

$$(x\cos\theta + y\sin\theta + d)(x\cos\varphi + y\sin\varphi + d^{'}) = 0 \tag{A.34}$$

(1) If $|\theta - \varphi| \neq 0$, π or 2π, the discriminant converges to

$$(x\sin(\theta-\varphi) - d^{'}\sin\theta + d\sin\varphi)^2 \geqslant 0 \text{ for any } x \in \mathbb{R} \tag{A.35}$$

The resulting equation (A.34) describes two intersecting lines.

(2) If $|\theta - \varphi| = 0$, π or 2π, we have following situations regarding Q_i:

[1] $Q_i \to Q \geqslant 0$ or $+\infty$. Then the solution set of (A.29) expands into the set of all real numbers. So the sequence of hyperbolas converges to a pair of parallel lines if $\theta - \varphi = \pi$ or a single line otherwise.

[2] $Q_i \to Q < 0$.

Apparently, $\sin\theta \sin\varphi$ is nonzero. T_i either has a limit $T < \infty$ or diverges to ∞. If $T_i \to T < \infty$, the numerator $d_i^{'} \sin\theta_i - d_i \sin\varphi_i \to 0$ since the denominator $\sin(\theta_i - \varphi_i) \to 0$. It follows that $d/d^{'} = \sin\varphi / \sin\theta > 0$. Since $\varphi - \theta = 0$ or 2π, $d = d^{'}$. Thus the limiting form of (A.34) is

$$(x\cos\theta + y\sin\theta + d)^2 = 0 \tag{A.36}$$

Furthermore, (A.29) has a solution set $x \in (-\infty, T - \sqrt{-Q}) \bigcup (T + \sqrt{-Q}, \infty)$. The above equation represents two opposite rays. See example (A.1).

However, if $T_i \to \infty$, (A.29) has a solution set of all real numbers. Therefore, (A.34) represents two coincident lines (single line) when $d^{'} = d$ and $|\theta - \varphi| = 0$ or 2π. Otherwise, it represents two parallel lines. See example (A.2) and (A.3).

[3] $Q_i \to -\infty$.

First, let us suppose $T_i \to T < \infty$. The solution set for (A.29) shrinks as $i \to \infty$ and becomes an empty set in the end. Geometrically, the hyperbolas escape from Ω, ending up with an empty set. But remember that the we must guarantees all model objects intersect with Ω.

Next, if $T_i \to \infty$, (A.29) has the following possible limiting solution sets:

(1) $\emptyset$ if $T_i - \sqrt{-Q_i} \to -\infty$ and $T_i + \sqrt{-Q_i} \to +\infty$.

(2) $(-\infty, T - \sqrt{-Q})$ or $(T + \sqrt{-Q}, +\infty)$ if one of $T_i - \sqrt{-Q_i}$ and $T_i + \sqrt{-Q_i}$ has a finite limit ($T - \sqrt{-Q}$ and $T + \sqrt{-Q}$ can not be both finite under the assumption). Without loss of generality, let us suppose $T_i - \sqrt{-Q_i} \to T - \sqrt{-Q} < \infty$. Then $\sin(\theta_i - \varphi_i)(T_i - \sqrt{-Q_i}) \to 0$. More precisely,

$$d_i^{'} \sin\theta_i - d_i \sin\varphi_i - \sqrt{-4a_i \sin\theta_i \sin\theta_i\varphi_i} \to d^{'} \sin\theta - d\sin\varphi = 0 \tag{A.37}$$

If $\theta - \varphi = 0$ or 2π and $d^{'} = d$. The limiting object represented by

$$(x\cos\theta + y\sin\theta + d)^2 = 0 \quad x \in (-\infty, T - \sqrt{-Q}) \tag{A.38}$$

corresponds to a ray. Otherwise, $d^{'} = d = 0$ and therefore the limiting object represented by

$$(x\cos\theta + y\sin\theta)^2 = 0 \quad x \in (-\infty, T - \sqrt{-Q}) \tag{A.39}$$

corresponds to a ray which lies on a straight line passing through the origin. See example (A.4).

(3) $(-\infty, +\infty)$ if either $T - \sqrt{-Q} \to +\infty$ or $T + \sqrt{-Q} \to -\infty$. We only need to study the case that $T - \sqrt{-Q} \to +\infty$. Since $d_i^{'}\sin\theta_i - d_i\sin\varphi_i \to c < \infty$, $d^{'} = d$ and $|\theta_i - \varphi_i| = 0$ or 2π if $c = 0$ and the limiting object corresponds to two coincident lines. If $c \neq 0$, $d^{'} \neq d$ and the limit object corresponds to two parallel lines. See example (A.5).

The following examples (A.1)-(A.3) are taken under the assumption $a_i \to 0$, $d_i^{'} \to d^{'} < \infty$, $\theta_i - \varphi_i \to 0, \pi$, or 2π and $Q_i \to Q < 0$.

EXAMPLE A.1. $\underline{T_i \to T < \infty}$. *Let us set* $\theta_i = \varphi_i + 1/i$, $\varphi = \pi/2$, $a_i = -1/i^2$ *and* $d_i^{'} = d_i + 1/i$. *As* $i \to \infty$, $T_i \to 1$ *and* $Q_i \to -4$. *Solving* (A.29) *gives* $x \in (-\infty, -1]\bigcup[3, \infty)$. *Then* (A.34) *has the form*

$$(y + d)^2 = 0 \quad x \in (-\infty, -1]\bigcup[3, \infty) \tag{A.40}$$

which represents two opposite rays. See Figure above.

EXAMPLE A.2. $\underline{T_i \to \infty}$. *Let* $\theta_i = \varphi_i + 1/i$, $\varphi = \pi/2$, $a_i = -1/i^2$ *and* $d_i^{'} = d_i + 1/\sqrt{i}$. *The limiting solution set for* (A.29) *contains all real numbers and* (A.34) *has the form*

$$(y + d)^2 = 0 \quad x \in (-\infty, \infty) \tag{A.41}$$

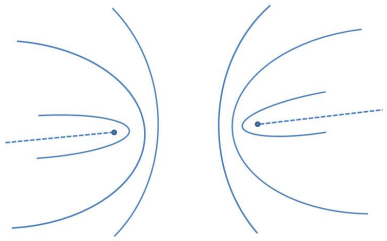


Figure A.8: Hyperbolas converge to two opposite rays

which represents two coincident lines.

EXAMPLE A.3. $\underline{T_i \to \infty}$. *Let* $\theta_i = \varphi_i + \pi + 1/i$, $\varphi = \pi/2$, $a_i = 1/i^2$ *and* $d_i^{'} = d_i + 1/\sqrt{i}$. (A.34) *has the form*

$$(y-d)(y+d) = 0 \quad x \in (-\infty, \infty) \tag{A.42}$$

which represents two parallel lines.

The following examples (A.4)-(A.5) are taken under the assumption $a_i \to 0$, $d_i^{'} \to d^{'} < \infty$, $\theta_i - \varphi_i \to 0, \pi$, or 2π and $Q_i \to -\infty$.

EXAMPLE A.4. $\underline{T_i \to \infty}$. *Let* $\theta_i = \varphi_i + 1/i$, $\varphi = \pi/2$, $a_i = -1/i$ *and* $d_i^{'} = d_i + 2/\sqrt{i}$. (A.34) *becomes*

$$(y+d)^2 = 0 \quad x \in [d, +\infty) \tag{A.43}$$

which represents a ray.

EXAMPLE A.5. $\underline{T_i \to \infty}$. *Let* $\theta_i = \varphi_i + 1/i$, $\varphi = \pi/2$, $a_i = -1/i$ *and* $d_i^{'} = d_i + 3/\sqrt{i}$. (A.34) *becomes*

$$(y+d)^2 = 0 \quad x \in (-\infty, +\infty) \tag{A.44}$$

which represents two coincident lines. For an example of two parallel lines, set $d_i^{'} = d_i + 1 + 3/\sqrt{i}$ *and keep the rest conditions.*

(II) $a_i \to a \neq 0$ or ∞.

(i) If $d_i \to d < \infty$ and $d_i^{'} \to \infty$, same strategy could be used from **(I) (i)**. The corresponding equation

$$(x\cos\theta_i + y\sin\theta_i + d_i)(\frac{x\cos\varphi_i + y\sin\varphi_i}{d_i^{'}} + 1) = \frac{a_i}{d_i^{'}} \tag{A.45}$$

transforms $x\cos\theta + y\sin\theta + d = 0$ corresponding to a straight line.

(ii) If $d_i \to \infty$ and $d_i^{'} \to \infty$, the situation becomes similar to **(I)(ii)**.

(iii) If $d_i \to d < \infty$ and $d_i^{'} \to d^{'} < \infty$, (A.26) transforms into

$$(x\cos\theta + y\sin\theta + d)(x\cos\varphi + y\sin\varphi + d^{'}) = a \tag{A.46}$$

(1) If $|\theta - \varphi| \neq 0$, π or 2π, the above equation represents a hyperbola.

(2) If $|\theta - \varphi| = 0$, π or 2π, let $x\cos\theta + y\sin\theta = t$. The resulting equation (A.46) becomes a quadratic equation in t:

$$t^2 + (d^{'} + d)t + dd^{'} - a = 0 \tag{A.47}$$

if $|\theta - \varphi| = 0$ or 2π or

$$t^2 + (d - d^{'})t - dd^{'} + a = 0 \tag{A.48}$$

if $|\theta - \varphi| = \pi$. Since $a \neq 0$, we only need to discuss the cases $Q_i \to \pm\infty$.

[1] $Q_i \to +\infty$. It is always possible to choose a coordinate system so that $\sin\theta\sin\varphi \neq 0$. Then the (A.29) has a solution set $x \in (-\infty, +\infty)$ for every i. Note that $a > 0$ when $|\theta - \varphi| = 0$, 2π and $a < 0$ when $|\theta - \varphi| = \pi$, leading to the positive discriminants for (A.47) and (A.48):

$$(d^{'} - d)^2 + 4a > 0 \qquad (d^{'} + d)^2 - 4a > 0 \tag{A.49}$$

Therefore, the resulting equation represents two parallel lines for both cases.

[2] $Q_i \to -\infty$. If $T_i \to T < \infty$, the same situation was explained in **(I)** **[3]** $Q_i \to -\infty$ and $T_i \to T < \infty$.

If $T_i \to \infty$, (A.29) has the following possible limiting solution sets:

(1) $\emptyset$ if $T_i - \sqrt{-Q_i} \to -\infty$ and $T_i + \sqrt{-Q_i} \to +\infty$.

(2) $(-\infty, T - \sqrt{-Q})$ or $(T + \sqrt{-Q}, +\infty)$ if one of $T_i - \sqrt{-Q_i}$ and $T_i + \sqrt{-Q_i}$ has a finite limit. Without loss of generality, let us suppose $T - \sqrt{-Q} < \infty$. Then

$$\sin(\theta_i - \varphi_i)(T_i - \sqrt{-Q_i}) \to d^{'} \sin\theta - d\sin\varphi - \sqrt{-4a\sin\theta\sin\varphi} = 0 \tag{A.50}$$

It follows that $(d^{'} - d)^2 + 4a = 0$ if $|\theta - \varphi| \to 0, 2\pi$, and $(d^{'} + d)^2 - 4a = 0$ if $|\theta - \varphi| = \pi$. Both of them ensure an unique solution for (A.47) and (A.48) respectively. So the hyperbola will converge to a ray in both cases.

(3) $(-\infty, \infty)$ if either $T_i - \sqrt{-Q_i} \to +\infty$ or $T_i + \sqrt{-Q_i} \to -\infty$. Without loss of generality, let us suppose $T_i - \sqrt{-Q_i} \to +\infty$. Then

$$\frac{d_i^{'}\sin\theta_i - d_i\sin\varphi_i}{\sin(\theta_i - \varphi_i)} > \sqrt{-\frac{4a_i\sin\theta_i\sin\varphi_i}{\sin^2(\theta_i - \varphi_i)}} \tag{A.51}$$

for sufficiently large $i's$. Squaring both sides yields

$$(d_i^{'}\sin\theta_i - d_i\sin\varphi_i)^2 > -4a_i\sin\theta_i\sin\varphi_i \tag{A.52}$$

which indicates that $(d^{'} - d)^2 + 4a > 0$ if $|\theta - \varphi| = 0$ or 2π or $(d^{'} + d)^2 - 4a > 0$ if $|\theta - \varphi| = \pi$. Thus both (A.47) and (A.48) represent a pair of parallel lines.

(III) $a_i \to \infty$.

(i) If $d_i \to d < \infty$ and $d_i^{'} \to d^{'} < \infty$, consider an equivalent form of (A.26)

$$\frac{(x\cos\theta_i + y\sin\theta_i + d_i)(x\cos\varphi_i + y\sin\varphi_i + d_i^{'})}{a_i} = 1 \tag{A.53}$$

The above equation becomes invalid when both sides reach at the limit. Same type of situation was explained in **(I)(ii)**.

(ii) If $d_i \to d < \infty$ and $d_i^{'} \to \infty$, let us consider the following possible limits of $\frac{d_i^{'}}{a_i}$.

(1) If $\frac{d_i^{'}}{a_i} \to 0$, (A.26) approximates to $0 = 1$. The equation of hyperbola becomes invalid, which indicates the violation of our assumption.

(2) If $\frac{d_i^{'}}{a_i} \to c \neq 0$ and ∞, (A.26) approximates an

$$x\cos\theta + y\sin\theta + d = \frac{1}{c} \tag{A.54}$$

as $i \to \infty$. We need to deal with following limits of Q_i.

[1] $Q_i \to +\infty$. Then (A.29) has a solution set of $x \in (-\infty.+\infty)$ for any sufficiently large i. So

$$x\cos\theta + y\sin\theta + d = \frac{1}{c} \quad x \in (-\infty. + \infty) \tag{A.55}$$

corresponds to a complete straight line.

[2] $Q_i \to -\infty$. Then (A.29) has a solution set $(-\infty, T_i - \sqrt{-Q_i}) \bigcup (T_i + \sqrt{-Q_i}, +\infty)$. Note that

$$\begin{aligned} T_i \pm \sqrt{-Q_i} &= \frac{d_i^{'}\sin\theta_i - d_i\sin\varphi_i}{\sin(\theta_i - \varphi_i)} \pm \sqrt{-\frac{4a_i\sin\theta_i\sin\varphi_i}{\sin^2(\theta_i - \varphi_i)}} \\ &= \frac{d_i^{'}}{\sin(\theta_i - \varphi_i)}(\sin\theta_i - \frac{d_i}{d_i^{'}}\sin\varphi_i \pm \sqrt{-\frac{4a_i\sin\theta_i\sin\varphi_i}{d_i^{'2}}}) \\ &\to +\infty \text{ or } -\infty \end{aligned} \tag{A.56}$$

So the solution set expands into $(-\infty, +\infty)$. The limit object corresponds to a straight line.

(3) $\frac{d_i^{'}}{a_i} \to \infty$

Same steps could be applied to show that $x \in (-\infty, +\infty)$ when the limit object is reached. So (A.26) approximates to

$$x\cos\theta + y\sin\theta + d = 0 \tag{A.57}$$

with $x \in (-\infty, +\infty)$ which corresponds to a straight line.

(iii) If $d_i \to \infty$ and $d_i^{'} \to \infty$, expand (A.26) into a standard form of the quadratic equation (see 4.1)

$$\begin{aligned}(\text{A.58}) \quad & x^2 \cos\theta_i \cos\varphi_i + xy(\cos\theta_i \sin\varphi_i + \sin\theta_i \cos\varphi_i) + y^2 \sin\theta_i \sin\varphi_i \\ & + x(d_i^{'} \cos\theta_i + d_i \cos\varphi_i) + y(d_i^{'} \sin\theta_i + d_i \sin\varphi_i) + d_i^{'} d_i - a_i = 0\end{aligned}$$

(1) $|\theta - \varphi| \neq \pi$. Note that

$$\begin{aligned}(\text{A.59}) \quad & (d_i^{'} \cos\theta_i + d_i \cos\varphi_i)^2 + (d_i^{'} \sin\theta_i + d_i \sin\varphi_i)^2 = d_i^{2'} + d_i^2 + 2 d_i d_i^{'} \cos(\theta_i - \varphi_i) \\ & \to +\infty\end{aligned}$$

So at least one of $d_i^{'} \cos\theta_i + d_i \cos\varphi_i$ and $d_i^{'} \sin\theta_i + d_i \sin\varphi_i$ must diverge to infinity. Without loss of generality, let us assume that

$$\left|\frac{d_i^{'} \cos\theta_i + d_i \cos\varphi_i}{d_i^{'} \sin\theta_i + d_i \sin\varphi_i}\right| \to |c| < +\infty$$

. Then

[1] If $\left|\frac{d_i^{'} d_i - a_i}{d_i^{'} \sin\theta_i + d_i \sin\varphi_i}\right| \to |c^{'}| < +\infty$, $d_i^{'} d_i - a_i \to \infty$ and dividing both sides of (A.58) by $d_i^{'} \sin\theta_i + d_i \sin\varphi_i$ yields

$$\begin{aligned}(\text{A.60}) \quad & \frac{x^2 \cos\theta_i \cos\varphi_i + xy(\cos\theta_i \sin\varphi_i + \sin\theta_i \cos\varphi_i) + y^2 \sin\theta_i \sin\varphi_i}{d_i^{'} \sin\theta_i + d_i \sin\varphi_i} \\ & + \frac{x(d_i^{'} \cos\theta_i + d_i \cos\varphi_i)}{d_i^{'} \sin\theta_i + d_i \sin\varphi_i} + y + \frac{d_i^{'} d_i - a_i}{d_i^{'} \sin\theta_i + d_i \sin\varphi_i} \to xc + y + c^{'} = 0\end{aligned}$$

which represents a straight line.

[2] If $\left|\frac{d_i^{'} d_i - a_i}{d_i^{'} \sin\theta_i + d_i \sin\varphi_i}\right| \to +\infty$, dividing both sides of (A.58) by $d_i^{'} d_i - a_i$ yields

$$\begin{aligned}(\text{A.61}) \quad & \frac{x^2 \cos\theta_i \cos\varphi_i + xy(\cos\theta_i \sin\varphi_i + \sin\theta_i \cos\varphi_i) + y^2 \sin\theta_i \sin\varphi_i}{d_i^{'} d_i - a_i} \\ & + \frac{x(d_i^{'} \cos\theta_i + d_i \cos\varphi_i)}{d_i^{'} d_i - a_i} + \frac{y(d_i^{'} \sin\theta_i + d_i \sin\varphi_i)}{d_i^{'} d_i - a_i} + 1 \to 1\end{aligned}$$

However, the left side of the equation stays at 0, leading to an impossible situation.

(2) $|\theta - \varphi| = \pi$. If one of last three coefficients: $d_i^{'} \cos\theta_i + d_i \cos\varphi_i$, $d_i^{'} \sin\theta_i + d_i \sin\varphi_i$ and $d_i^{'} d_i - a_i$ diverge to infinity, we can follow the similar steps as above in **(1)** and

draw the same conclusion. So we skip the technical detail and proceed to the case when all three coefficients have finite limits. Recall that a quadratic curve in defined by an equation

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0 \tag{A.62}$$

in section (4.2). Note that $J = B^2 - AC = \cos(\theta_i - \varphi_i) \to 0$ as $|\theta_i - \varphi_i| \to \pi$. So the limit object could be a parabola, two parallel lines or coincident line depending on the value of Δ and K (see section (4.2)) .

$$\begin{aligned}\Delta = &-\frac{\sin^2(\theta_i - \varphi_i)(d_i^{'}d_i - a_i)}{4} - \frac{\cos\theta_i \cos\varphi_i(d_i^{'}\sin\theta_i + d_i\sin\varphi_i)^2}{4} \\ &+ \frac{\sin(\theta_i + \varphi_i)(d_i^{'}\sin\theta_i + d_i\sin\varphi_i)(d_i^{'}\cos\theta_i + d_i\cos\varphi_i)}{4} \\ &- \frac{\sin\theta_i \sin\varphi_i(d_i^{'}\cos\theta_i + d_i\cos\varphi_i)^2}{4}\end{aligned} \tag{A.63}$$

The first term goes vanish since $d_i^{'}d_i - a_i$ has a finite limit. So we only need to look at the remaining terms in the expression. Without loss of generality, Suppose $|d_i^{'}/d_i| = c < \infty$. So

$$\begin{aligned}d_i^2(&-\frac{\cos\theta_i \cos\varphi_i(d_i^{'}/d_i\sin\theta_i + \sin\varphi_i)^2}{4} \\ &+ \frac{\sin(\theta_i + \varphi_i)(d_i^{'}/d_i\sin\theta_i + \sin\varphi_i)(d_i^{'}/d_i\cos\theta_i + \cos\varphi_i)}{4} \\ &- \frac{\sin\theta_i \sin\varphi_i(d_i^{'}/d_i\cos\theta_i + \cos\varphi_i)^2}{4}) = d_i^2\Delta^{'}\end{aligned} \tag{A.64}$$

which has the same finite limit as Δ. Note that $\Delta^{'} \to 0$ while $d_i^2 \to +\infty$. So if $d_i^2\Delta^{'} \to c^{'} \neq 0$, the limit object will be a parabola.

However, if $d_i^2\Delta^{'} \to 0$, the type of limit object is determined by

$$\begin{aligned}K &= (A + C)F - D^2 - E^2 \\ &= \cos(\theta_i - \varphi_i)(d_i d_i^{'} - a_i) - \frac{d_i^2 + d_i^{'2} + 2d_i d_i^{'}\cos(\theta_i - \varphi_i)}{4}\end{aligned} \tag{A.65}$$

(A.66)

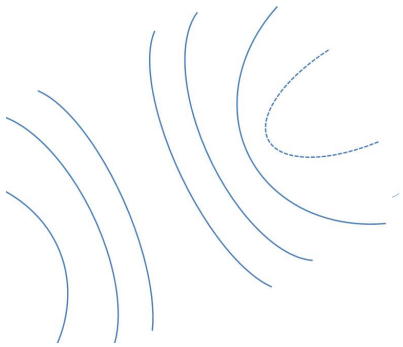


Figure A.9: Hyperbolas to Parabola

Type of curve	limit point(s)
pairs of opposite rays	ray, line
intersecting lines	two coincident lines, two parallel lines
ray	line

which is negative when (A.62) represents a hyperbola. Then we have two types of limit objects based on the its limit value: a pair of parallel lines if $K \to K_0 < 0$ or two coincident lines if $K = 0$ (Also $D, E \to 0$). See example (A.6).

The following examples (A.6)- are taken under the assumption $a_i \to \infty$, $d_i \to \infty$, $d_i' \to \infty$, $\theta_i - \varphi_i \to \pi$.

EXAMPLE A.6. *Let* $\theta_i = \frac{\pi}{2} + 1/i$, $\varphi_i = \frac{3\pi}{2} + \frac{1}{i}$, $d_i = d_i' = i$ *and* $a_i = i^2$. *Then* $J = \frac{\sin(1/i)}{2} \to 0$ *and* $\Delta = i^2 \frac{\sin^2(1/i)}{4} \to 1/4$. *So the limiting object will be a parabola. But if* $\Delta = i^2 \frac{\sin^2(1/i)}{4} \to 1/4$. $d_i = d_i' = \sqrt{i}$, $\Delta = i \frac{\sin^2(1/i)}{4} \to 0$ *and* $K = -\frac{2i^2 - 2i^2 \cos(\pi + 1/i)}{4} \to 0$. *The limit object can be identified as two coincident lines based on our table in section* (4.2).

So now we finished our search for the limit point of a sequence of hyperbolas. As for a sequence of pairs of opposite rays, intersecting lines and single rays , we skipped the annoying technical steps and provides their limits below:

It is unnecessary to repeat the search for the limit points of sequences of other types of curves since these have been discussed in the proof of theorem (A.4). So theorem (A.5) is justified. □

$$\text{dist}(P, S) = \min_{Q \in S} \text{dist}(P, Q)$$

Note that the minimum can be always achieved on S if S is closed.

Proof: Let $\{Q_i\}$ be a sequence of points such that $\text{dist}(P, Q_i) \to \inf_{Q \in S} \text{dist}(P, Q)$ as $i \to \infty$. Suppose $\text{dist}(P, Q_i) < r$ for $i = 1, 2, ...$ Then the closed disk $B(P, r) = \{x | \text{dist}(P, x) \leqslant r\}$ contains the sequence $\{Q_i\}$. Furthermore, $B(P, r) \bigcap S$ is a compact set in R^2. So there exists a convergent subsequence $\{Q_{i_j}\}$ with a limit $Q_0 \in B(P, r) \bigcap S$. Since $\lim_{i \to \infty} \text{dist}(P, Q_{i_j}) = \text{dist}(P, Q_0) = \inf_{Q \in S} \text{dist}(P, Q)$. The minimum is attained by $Q_0 \in S$.

A.3. Upper bounds for the partial derivatives

In this section we will derive the formulas for the first and the second partial derivatives for a signed distance(i.e the actual distance combined with a sign depending on whether the point is inside or outside the ellipse) between the point and the ellipse with respect to each geometric parameters: the major axis a, minor axis b, coordinates of the center (c_1, c_2) and the angle θ between the major axis and the x axis and then determine their upper bounds. Before running into the technical detail, we'll provide a glimpse of our final results:

Let (x, y) be the given point, (u, v) its projection on a given conic (ellipse, hyperbola, parabola) and Θ the parameter vector for the conic which can be described as

$$P(u, v; \Theta) = 0 \tag{A.67}$$

Furthermore, let us denoted by d the distance between the given point and the ellipse. If α and β are two independent parameters for the conic, then we have

$$d_\alpha = \frac{P_\alpha}{\sqrt{P_u^2 + P_v^2}}$$

(see [**16**]) and

(A.68)

$$d_\alpha d_\beta + d d_{\alpha\beta} = tP_{\alpha\beta} - (P_\alpha, tP_{u\alpha}, tP_{v\alpha}) \begin{pmatrix} 0 & P_u & P_v \\ P_u & 1+tP_{uu} & tP_{uv} \\ P_v & tP_{uv} & 1+tP_{vv} \end{pmatrix}^{-1} \begin{pmatrix} P_\beta \\ tP_{u\beta} \\ tP_{v\beta} \end{pmatrix}$$

where $t = \frac{d}{\sqrt{P_u^2+P_v^2}}$. The upper bounds for partial derivatives with respect to each geometric parameters are:

(1)major axis a:

$$|d_a| \leq 1$$

(2)minor axis b

$$|d_b| \leq 1$$

(3)center coordinates c_1 and c_2

$$d_{c_1}^2 + d_{c_2}^2 = 1$$

(4)angle θ

$$|d_\theta| \leq c = \sqrt{a^2 + b^2}$$

Technical calculations

By orthogonal conditions, we have

(A.69) $$x - u = t \cdot P_u \quad \text{and} \quad y - v = t \cdot P_v$$

where t is a parameter depending on the distance d. The square of the distance between the points and the conic is

(A.70) $$d^2 = (x-u)^2 + (y-v)^2$$

Differentiating (A.67),(A.69) and (A.70), we have

$$P_\alpha = -P_u u_\alpha - P_v v_\alpha \tag{A.71}$$

$$-u_\alpha = t_\alpha P_u + t(P_u)_\alpha \qquad -v_\alpha = t_\alpha P_v + t(P_v)_\alpha \tag{A.72}$$

$$d \cdot (d)_\alpha = -(x-u)u_\alpha - (y-v)v_\alpha \tag{A.73}$$

Substituting (A.72) into (A.71),

$$\begin{aligned} P_\alpha &= P_u(t_\alpha P_u + t(P_u)_\alpha) + P_v \cdot (t_\alpha P_v + t(P_v)_\alpha) \\ &= t_\alpha(P_u^2 + P_v^2) + t(P_u(P_u)_\alpha + P_v(P_v)_\alpha) \end{aligned} \tag{A.74}$$

In (A.73), we can replace u_α and v_α by (A.72), $(x-u)$ and $(y-v)$ by (A.69):

$$\begin{aligned} d \cdot d_\alpha &= -(x-u)u_\alpha - (y-v)v_\alpha \\ &= tP_u \cdot (t_\alpha P_u + t(P_u) + tP_v \cdot (t_\alpha P_v + t(P_v) \\ &= t \cdot t_\alpha(P_u^2 + P_v^2) + t^2 \cdot (P_u(P_u)_\alpha + P_v(P_v)_\alpha) \\ &= t \cdot [t_\alpha(P_u^2 + P_v^2) + t(P_u(P_u)_\alpha + P_v(P_v)_\alpha)] \\ &= t \cdot P_\alpha \end{aligned} \tag{A.75}$$

The last equality follows from (A.74). If we replace $x-u$ and $u-v$ in (A.70) by (A.69), we will have the following:

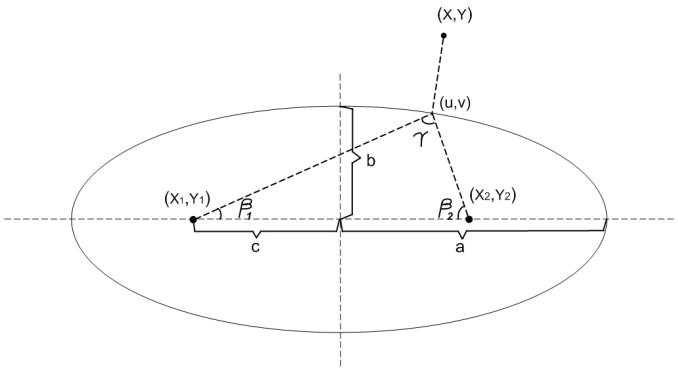
$$d^2 = t^2(P_u^2 + P_v^2)$$

Thus yielding,

$$d = t \cdot \sqrt{P_u^2 + P_v^2} \tag{A.76}$$

It follows that

$$\begin{aligned} d \cdot d_\alpha &= t \cdot P_\alpha \\ &= \frac{d}{\sqrt{P_u^2 + P_v^2}} \cdot P_\alpha \end{aligned} \tag{A.77}$$



Finally, we obtain a simple form for d_α:

$$d_\alpha = \frac{P_\alpha}{\sqrt{P_u^2 + P_v^2}} \tag{A.78}$$

Let (x_1, y_1) and (x_2, y_2) be the two focuses of a ellipse and a be the semi-major axis. Then the ellipse is described by the equation:

$$\sqrt{(u-x_1)^2 + (v-y_1)^2} + \sqrt{(u-x_2)^2 + (v-y_2)^2} - 2a = 0 \tag{A.79}$$

By differentiating (A.79) with respect to u and v, we have

$$P_u = ((u-x_1)^2 + (v-y_1)^2)^{-1/2}(u-x_1) + ((u-x_2)^2 + (v-y_2)^2)^{-1/2}(u-x_2) \tag{A.80}$$

$$P_v = ((u-x_1)^2 + (v-y_1)^2)^{-1/2}(v-y_1) + ((u-x_2)^2 + (v-y_2)^2)^{-1/2}(v-y_2) \tag{A.81}$$

So

$$\sqrt{P_u^2 + P_v^2} = \sqrt{2 + 2\frac{(u-x_1)(u-x_2) + (v-y_1)(v-y_2)}{\sqrt{(u-x_1)^2 + (v-y_1)^2}\sqrt{(u-x_2)^2 + (v-y_2)^2}}} \tag{A.82}$$

$$= \sqrt{2(1 + cos\gamma)} \tag{A.83}$$

where γ is the angle between two rays joining the point (u, v) and two focuses(see illustration in the figure above).

To rewrite (A.79) in geometric parameters (center (c_1, c_2), major axis a, minor axis b and the angle θ between major axis and the x axis), we use following relations:

$$x_1 = c_1 + \sqrt{a^2 - b^2}cos\theta \qquad x_2 = c_1 - \sqrt{a^2 - b^2}cos\theta$$
$$y_1 = c_2 + \sqrt{a^2 - b^2}cos\theta \qquad y_2 = c_2 - \sqrt{a^2 - b^2}cos\theta$$

For the purpose of simplicity , we will use following shorthand notations:

$$d_{f1} = \sqrt{(u - x_1)^2 + (v - y_1)^2}$$
$$d_{f2} = \sqrt{(u - x_2)^2 + (v - y_2)^2}$$
$$c = \sqrt{a^2 - b^2}$$

Then we have derivatives with respect to each geometric parameters and their upper bounds:

(i) $|d_a| \leq 1$

$$d_a = \frac{P_a}{\sqrt{P_u^2 + P_v^2}} = \frac{\frac{a}{c}(-\frac{(u-x_1)cos\theta}{d_{f1}} - \frac{(v-y_1)sin\theta}{d_{f1}} + \frac{(u-x_2)cos\theta}{d_{f2}} + \frac{(v-y_2)sin\theta}{d_{f2}}) - 2}{\sqrt{2(1 + cos\gamma)}} \tag{A.84}$$

Let $s = \frac{(2a+2c)}{2} = a + c$. Then

$$cos\frac{\gamma}{2} = \sqrt{\frac{s(s - 2c)}{d_{f1}d_{f2}}} = \sqrt{\frac{a^2 - c^2}{d_{f1}d_{f2}}}$$

$$d_{f1}d_{f2} = \frac{a^2 - c^2}{cos^2\frac{\gamma}{2}}$$

we will use them later for further transformations.

Since the upper bound of the partial derivatives does not depend on θ, (A.84) can be simplified by setting $\theta = 0$. Then

$$\begin{aligned}
d_a &= \frac{\frac{a}{c}(cos\beta_1 + cos\beta_2) - 2}{\sqrt{2(1+cos\gamma)}} \\
&= \frac{\frac{a}{c}(\frac{d_{f1}^2+4c^2-d_{f2}^2}{4d_{f1}c} + \frac{d_{f2}^2+4c^2-d_{f1}^2}{4d_{f2}c}) - 2}{\sqrt{2(1+cos\gamma)}} \\
&= \frac{\frac{a}{c}(\frac{d_{f1}+d_{f2}}{4c} + \frac{c(d_{f1}+d_{f2})}{d_{f1}d_{f2}} - \frac{d_{f1}^3+d_{f2}^3}{4cd_{f1}d_{f2}}) - 2}{\sqrt{2(1+cos\gamma)}} \\
&= \frac{\frac{a}{c}(\frac{a}{2c} + \frac{2ac\cdot cos^2\frac{\gamma}{2}}{a^2-c^2} - \frac{(d_{f1}+d_{f2})((d_{f1}+d_{f2})^2-3d_{f1}d_{f2})}{4cd_{f1}d_{f2}}) - 2}{\sqrt{2(1+cos\gamma)}} \\
&= \frac{\frac{a}{c}(\frac{a}{2c} + \frac{2ac\cdot cos^2\frac{\gamma}{2}}{a^2-c^2} - \frac{8a^3cos^2\frac{\gamma}{2}}{4c(a^2-c^2)} + \frac{3a}{2c}) - 2}{\sqrt{2(1+cos\gamma)}} \\
&= \frac{\frac{a}{c}(\frac{2a}{c} - \frac{2a}{c}cos^2\frac{\gamma}{2}) - 2}{2cos\frac{\gamma}{2}} \\
&= \frac{\frac{a^2}{c^2}(1-cos^2\frac{\gamma}{2}) - 1}{2cos\frac{\gamma}{2}}
\end{aligned}$$

Note that $cos\frac{\gamma}{2} \in [\frac{b}{a}, 1]$. Set $x = cos\frac{\gamma}{2}$. Then

$$F(x) = \frac{\frac{a^2}{c^2}(1-x^2) - 1}{x} = -\frac{a^2}{c^2}x + + \frac{\frac{a^2}{c^2} - 1}{x}$$

is monotonically decreasing on $[\frac{b}{a}, 1](F'(x) < 0)$. Hence,$\mid d_a \mid \leq \mid F(1) \mid = 1$.

(ii)$|d_b| \leq 1$

$$d_b = \frac{P_b}{\sqrt{P_u^2 + P_v^2}} \qquad = \frac{\frac{b}{c}(\frac{(u-x_1)cos\theta}{d_{f1}} + \frac{(v-y_1)sin\theta}{d_{f1}} - \frac{(u-x_2)cos\theta}{d_{f2}} - \frac{(v-y_2)sin\theta}{d_{f2}})}{\sqrt{2(1+cos\gamma)}} \tag{A.85}$$

By taking advantage of similarity between (A.84) and (A.85), we can skip intermediate steps and reach the last equality:

$$d_b = \frac{\frac{ab(cos^2\frac{\gamma}{2}-1)}{c^2}}{cos\frac{\gamma}{2}}$$

Let $x = cos\frac{\gamma}{2}$. Then $g(x) = \frac{ab(x-\frac{1}{x})}{c^2}$ is increasing on $[\frac{b}{a}, 1]$. So $\mid d_b \mid \leq 1$.

(iii)$|d_\theta| \leq c$

Since the derivative is not affected by the value of θ, let us set $\theta = \frac{\pi}{2}$.Then

$$\begin{aligned}
d_\theta|_{\theta=\pi/2} &= \frac{c(\frac{u-x_1}{d_{f1}} - \frac{u-x_2}{d_{f2}})}{2cos\frac{\gamma}{2}} \\
&= \frac{c(-cos\beta_1 + cos\beta_2)}{2cos\frac{\gamma}{2}} \\
&= \frac{c \cdot (-\frac{d_{f1}^2+4c^2-d_{f2}^2}{4d_{f1}c} + \frac{d_{f2}^2+4c^2-d_{f1}^2}{4d_{f2}c})}{2cos\frac{\gamma}{2}} \\
&= \frac{d_{f1}d_{f2}^2 - d_{f1}^2d_{f2} + 4c^2(d_{f1} - d_{f2}) - d_{f1}^3 + d_{f2}^3}{8d_{f1}d_{f2}cos\frac{\gamma}{2}} \\
&= \frac{(d_{f1} - d_{f2})(4c^2 - d_{f1}d_{f2} - (d_{f1} + d_{f2})^2 - d_{f1}d_{f2}))}{8d_{f1}d_{f2}cos\frac{\gamma}{2}} \\
&= \frac{(d_{f1} - d_{f2})(4c^2 - 4a^2)}{8d_{f1}d_{f2}cos\frac{\gamma}{2}} \\
&= -\frac{(d_{f1} - d_{f2})cos^2\frac{\gamma}{2}}{2cos\frac{\gamma}{2}} \\
&= -\frac{1}{2}(d_{f1} - d_{f2})cos\frac{\gamma}{2}
\end{aligned}$$

So $\mid d_\theta \mid = \mid \frac{1}{2}(d_{f1} - d_{f2})cos\frac{\gamma}{2} \mid \leq c.$

(iv)$d_{c_1}^2 + d_{c_2}^2 = 1$

$$\begin{aligned}
d_{c_1} &= -\frac{\frac{u-x_1}{d_{f1}} + \frac{u-x_2}{d_{f2}}}{\sqrt{2(1+cos\gamma)}} \\
d_{c_2} &= -\frac{\frac{v-y_1}{d_{f1}} + \frac{v-y_2}{d_{f2}}}{\sqrt{2(1+cos\gamma)}}
\end{aligned}$$

It follows that

$$\begin{aligned}
d_{c_1}^2 + d_{c_2}^2 &= \frac{1 + 1 + 2\frac{(u-x_1)(u-x_2)+(v-y_1)(v-y_2)}{d_{f1}d_{f2}}}{2(1+cos\gamma)} \\
&= \frac{2(1+cos\gamma)}{2(1+cos\gamma)} \\
&= 1
\end{aligned}$$

Next we will derive the formula for the second derivatives of the objective function $\mathcal{F}$. Notice that in (A.72) and (A.74)

$$(P_u)_\alpha = P_{uu}u_\alpha + P_{uv}v_\alpha + P_{u\alpha}$$

$$(P_v)_\alpha = P_{vv}v_\alpha + P_{uv}u_\alpha + P_{v\alpha}$$

By substitution, they turn into

$$t_\alpha P_u + (1 + tP_{uu})u_\alpha + tP_{uv}v_\alpha = -tP_{u\alpha} \tag{A.86}$$

$$t_\alpha P_v + tP_{uv}v_\alpha + (1 + tP_{vv})u_\alpha = -tP_{v\alpha} \tag{A.87}$$

$$t_\alpha(P_u^2+P_v^2)+t(P_uP_{uu}+P_vP_{uv})u_\alpha+t(P_uP_{uv}+P_vP_{uv})v_\alpha = P_\alpha-t(P_uP_{u\alpha}+P_vP_{v\alpha}) \tag{A.88}$$

With (A.86)$\cdot P_u$+(A.87)$\cdot P_v$- (A.88), we have

$$P_u u_\alpha + P_v v_\alpha = -P_\alpha \tag{A.89}$$

Then

$$\begin{pmatrix} 0 & P_u & P_v \\ P_u & 1+tP_{uu} & tP_{uv} \\ P_v & tP_{uv} & 1+tP_{vv} \end{pmatrix} \cdot \begin{pmatrix} t_\alpha \\ u_\alpha \\ v_\alpha \end{pmatrix} = \begin{pmatrix} -P_\alpha \\ -tP_{u\alpha} \\ -tP_{v\alpha} \end{pmatrix}$$

$$\begin{pmatrix} t_\alpha \\ u_\alpha \\ v_\alpha \end{pmatrix} = -\begin{pmatrix} 0 & P_u & P_v \\ P_u & 1+tP_{uu} & tP_{uv} \\ P_v & tP_{uv} & 1+tP_{vv} \end{pmatrix}^{-1} \cdot \begin{pmatrix} P_\alpha \\ tP_{u\alpha} \\ tP_{v\alpha} \end{pmatrix}$$

Differentiating (A.75) with respect to some parameter β yielding

$$\begin{aligned} d_\alpha d_\beta + d d_{\alpha\beta} &= t_\beta P_\alpha + t(P_{u\alpha} u_\beta + P_{v\alpha} v_\beta + P_{\alpha\beta}) \\ &= (P_\alpha, tP_{u\alpha}, tP_{v\alpha}) \cdot \begin{pmatrix} t_\beta \\ u_\beta \\ v_\beta \end{pmatrix} + tP_{\alpha\beta} \\ &= tP_{\alpha\beta} - (P_\alpha, tP_{u\alpha}, tP_{v\alpha}) \begin{pmatrix} 0 & P_u & P_v \\ P_u & 1+tP_{uu} & tP_{uv} \\ P_v & tP_{uv} & 1+tP_{vv} \end{pmatrix}^{-1} \begin{pmatrix} P_\beta \\ tP_{u\beta} \\ tP_{v\beta} \end{pmatrix} \end{aligned}$$

Therefore,

$$\begin{aligned} \mathcal{F}_{\alpha\beta} &= 2\sum (d_\alpha d_\beta + d d_{\alpha\beta}) \\ &= 2\sum [tP_{\alpha\beta} - (P_\alpha, tP_{u\alpha}, tP_{v\alpha}) \begin{pmatrix} 0 & P_u & P_v \\ P_u & 1+tP_{uu} & tP_{uv} \\ P_v & tP_{uv} & 1+tP_{vv} \end{pmatrix}^{-1} \begin{pmatrix} P_\beta \\ tP_{u\beta} \\ tP_{v\beta} \end{pmatrix}] \end{aligned}$$

A.4. Fine structure of parameter space

LEMMA A.4. *Let $P = (x_0, y_0) \in S_0$, there exists a point (x, y) on S such that* $\mathrm{dist}((x_0, y_0), (x, y)) < \varepsilon$ *if* $\mathrm{dist}(\mathbf{P}, \mathbf{P}_0) < \delta$ *for some* $\delta > 0$.

PROOF. First, let S_0 be a real conic other than single point or two coincident lines. Then by Lemma A.5, for any point $(x_0, y_0) \in S_0$ satisfying

$$F(x, y|\mathbf{P}_0) = A_0 x_0^2 + 2B_0 x_0 y_0 + C_0 y_0^2 + 2D_0 x_0 + 2E_0 y_0 + F_0 = 0, \tag{A.90}$$

there exists two points (x_-, y_-) and (x_+, y_+) within the $\varepsilon-$ open neighborhood of (x_0, y_0) denoted by $U((x_0, y_0), \varepsilon)$ such that $F(x_-, y_-|\mathbf{P}_0) < 0$ and $F(x_+, y_+|\mathbf{P}_0) > 0$. Note that

$$G(\mathbf{P}|x, y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F$$

is a continuous function on S^5. So $G(\mathbf{P}|x_-, y_-) = F(x_-, y_-|\mathbf{P}) < 0$ and $G(\mathbf{P}|x_+, y_+) = F(x_+, y_+|\mathbf{P}) > 0$ if $\text{dist}(\mathbf{P}, \mathbf{P}_0) < \delta$ for some $\delta > 0$. Then by intermediate value theorem, there exists a point $(x, y) \in U((x_0, y_0), \varepsilon)$ such that $F(x, y|\mathbf{P}) = 0$, which partially completes the proof. Next, let us check the case that S_0 is a single point in R^2.

Single point

A single point P with coordinates (x_0, y_0) is defined by equation

$$a^2(x - x_0)^2 + b^2(y - y_0)^2 + 2c(x - x_0)(y - y_0) = 0$$

where a and b are arbitrary non-zero numbers and $c^2 < a^2b^2$ (see section 4.2).

For any $\mathbf{P}_0 \in S^5$ corresponding to a single point (x_0, y_0), there exists an open neigborhood U containing $\mathbf{P}_0$ such that $\mathbf{P} \in U$ either corresponds to a single point or an ellipse (see Figure 4.4). Also note that the quadratic equation for a given ellipse takes the following form

$$d(x - x_0^{'})^2 + f(y - y_0^{'})^2 + 2e(x - x_0^{'})(y - y_0^{'}) = l$$

where $l \neq 0$. The ellipse has its center Q at $(x_0^{'}, y_0^{'})$, and length of major axis a:

$$a = \frac{2\sqrt{2l}}{\sqrt{(d + f) - \sqrt{(d - f)^2 + 4e^2}}}$$

Since

$$Q = (x_0^{'}, y_0^{'}) \to P = (x_0, y_0)$$

$$\sqrt{(d^2 + f^2) - \sqrt{(d^2 - f^2)^2 + 4e^2}} \to \sqrt{(a^2 + b^2) - \sqrt{(a^2 - b^2)^2 + 4c^2}} > 0$$

$$d \to 0$$

as $\mathbf{P} \to \mathbf{P}_0$, $a < \varepsilon$ and $\text{dist}(P, Q) < \varepsilon$ if $\text{dist}(\mathbf{P}, \mathbf{P}_0) < \delta$. So if $\mathbf{P}$ corresponds to a single point (when $d = 0$) which is $Q = (x_0^{'}, y_0^{'})$, $\text{dist}(P, Q) < \varepsilon$. Or if $\mathbf{P}$ corresponds to an ellipse ($d \neq 0$) with its center Q and major axis $a < \varepsilon$, there exists a point P' on that ellipse such that $\text{dist}(P, P') < \varepsilon$. This completes the proof of lemma (A.4). $\square$

LEMMA A.5. *For any point (x_0, y_0) on a given model object S_0 (other than single point or coincident lines) described by the general quadratic equation*

$$F(x, y|\mathbf{P}_0) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

where $\mathbf{P}_0 = (A, B, C, D, E, F) \in S^5$ and $\varepsilon > 0$, there exists two points (x_-, y_-) and (x_+, y_+) belongs to the ε neigborhood of (x, y) denoted by $U((x, y), \varepsilon)$ such that $F(x_-, y_-|\mathbf{P}_0) < 0$ and $F(x_+, y_+|\mathbf{P}_0) > 0$.

PROOF. Let S_0 be following model objects:

(i) **Hyperbola, parabola, ellipse and a pair of parallel lines**

For a given point $(x_0, y_0) \in S_0$, pick $(x_0^{'}, y_0^{'})$ $(x_0 \neq x_0^{'})$ on S_0. Then there exists a straight line represented by a linear equation $y = ax + b$ passing through both points. By substitution, (A.94) turns into a quadratic equation $F(x, ax + b) = 0$ which has two distinct solutions x_0 and x_0'. Since $F(x, ax + b) = 0$ is strictly monotonic around x_0, there exists x_- and x_+ within a ε neighborhood of x_0 such that $F(x_+, ax_+ + b) > 0$ and $F(x_-, ax_- + b) < 0$. Set $y_* = ax_* + b$ $(* = +, -)$. Therefore, $F(x_-, y_-|\mathbf{P}_0) < 0$ and $F(x_+, y_+|\mathbf{P}_0) > 0$. for (x_-, y_-) and $(x_+, y_+) \in U((x_0, y_0), \varepsilon)$.

(ii) **Straight line**

Any straight line can be expressed in a linear equation

$$F(x, y) = 2Dx + 2Ey + F = 0$$

where $D^2 + E^2 \neq 0$. Let us assume $D > 0$(any other different conditions can be verified in a similar manner). Since $2Dx_0 + 2Ey_0 + F = 0$, $F(x_0 + \varepsilon/2, y_0) = D\varepsilon > 0$ and $F(x_0 - \varepsilon/2, y_0) = -D\varepsilon < 0$. So the conclusion is justified for straight lines.

(iii) **Intersecting lines**

The quadratic equation $F(x, y)$ for a pair of intersecting lines can be transformed into a product of two linear relations:

$$(2Dx + 2Ey + F)(2D'x + 2E'y + F') = 0$$

where each linear relation on the left corresponds to one of straight lines. For every point (x_0, y_0) satisfying $2Dx_0 + 2Ey_0 + F = 0$ and $2D^{'}x_0 + 2E^{'}y_0 + F > 0$ (the proofs for other conditions are similar), it belongs to one line but not on the other. By the proof showed in (ii), there exists two points (x_-, y_-) and (x_+, y_+) within $U((x_0, y_0), \varepsilon)$ for any $\varepsilon > 0$, such that $2Dx_- + 2Ey_- + F < 0$ and $2Dx_+ + 2Ey_+ + F > 0$. Note that $2D^{'}x_* + 2E^{'}y_* + F^{'} > 0$ $(* = +, -)$ for sufficiently small ε. Then

$$(2Dx_+ + 2Ey_+ + F)(2D^{'}x_+ + 2E^{'}y_+ + F^{'}) > 0$$

$$(2Dx_- + 2Ey_- + F)(2D^{'}x_- + 2E^{'}y_- + F^{'}) < 0$$

If (x_0, y_0) is at the intersection of two lines so that both factor the left side of (A.4) vanish at (x_0, y_0), pick a point $(x_0^{'}, y_0^{'}) \in U((x_0, y_0), \varepsilon/2)$ on one of the straight line but on the other. Then by the proof of the first part, there exists two point (x_-, y_-) and (x_+, y_+) within $U((x_0^{'}, y_0^{'}), \varepsilon/2)$ and certainly within $U((x_0, y_0), \varepsilon)$ such that $F(x_-, y_-) < 0$ and $F(x_+, y_+) > 0$. The proof of lemma (A.5) is completed. □

THEOREM A.6. *(**Convergence of conics: general case**) Suppose a sequence* $\mathbf{P}_i$ *of parameter vectors corresponding to real (not imaginary) conics,* S_n*, converges to a parameter vector* $\mathbf{P}$ *corresponding to a real (not imaginary) conic,* S*, which is not a pair of coincident lines, i.e.,* $\mathbf{P} \notin \mathbb{D}_{\mathrm{CL}}$*. Then* $S_n \to S$ *geometrically, in the sense of section* (2.2)*.*

PROOF. Let R be a rectangle in $\mathbb{R}^2$:

$$R = \{-A \le x \le A, \ \ -B \le y \le B\}, \tag{A.91}$$

Let's consider following conditions:

(i)$S \cap R = \emptyset$

Since R is a compact set in $\mathbb{R}^2$, the continuous function

$$F(x, y|\mathbf{P}) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F > M > 0 \quad (\text{or} < M < 0) \tag{A.92}$$

for all $(x, y) \in R$. Then $F(x, y|\mathbf{P}_i) > M - \varepsilon > 0$ for $|\mathbf{P}_i - \mathbf{P}| < \delta$. By the assumption of convergence of $\mathbf{P}_i$, there exists a j such that $|\mathbf{P}_i - \mathbf{P}| < \delta$ for $i > j$. Thus S_i does not intersect with R and

$$\mathrm{dist}_{\mathrm{H}}(S_i, S; R) = 0$$

for $i > j$. Now we have

$$\mathrm{dist}_{\mathrm{H}}(S_i, S; R) \to 0 \quad \text{as} \quad i \to \infty. \tag{A.93}$$

(ii)$S \cap R \neq \emptyset$

Let $B(P, r)$ be the open disk of radius r with its center $P \in \mathbb{R}^2$. Then for any $\varepsilon > 0$ and the set $Q = \{P|P \in S \cap R\}$ there exists finitely many distinct open discs $B(P_k, \varepsilon/2)$ $(k = 1, 2, \cdots, n)$ with $P_k \in Q$ such that $Q \subset \bigcup_{k=1}^{n} B(P_k, \varepsilon/2)$. So by lemma (A.4), for each P_k, there exists a $P_k^{'} \in S_i$ (associated with the parameter vector $\mathbf{P}_i$) such that $\mathrm{dist}(P_k^{'}, P_k) < \varepsilon/2$ for $|\mathbf{P}_i - \mathbf{P}| < \delta_k$. Apparently, $\mathrm{dist}(P_k^{'}, P) < \varepsilon$ if $P \in B(P_k, \varepsilon/2)$. Set $\delta = \min(\delta_k)$. Since $\mathbf{P}_i$ converges to $\mathbf{P}$, $|\mathbf{P}_i - \mathbf{P}| < \delta$ when $i > N$. Then $\mathrm{dist}(P, S_i) < \varepsilon$ for $P \in S \cap R$ if $i > N$, implying $\sup_{P \in S \cap R} \mathrm{dist}(P, S_i) < \varepsilon$. Hence

$$\sup_{P \in S \cap R} \mathrm{dist}(P, S_i) \to 0$$

Next, let us show

$$\sup_{Q \in S_i \cap R} \mathrm{dist}(Q, S) \to 0$$

Suppose above condition does not hold. By way of contradiction there exists a $\sigma > 0$ such that for any $N > 0$ one can always find some $i_0 > N$ with $\sup_{Q \in S_{i_0} \cap R} \mathrm{dist}(Q, S) > \sigma$. So $\mathrm{dist}(Q, S) > \sigma/2$ for some $Q \in S_{i_0} \cap R$. In other words, we can find a sequence of conics S_{i_j} each containing a point $Q_{i_j} \in S_{i_j} \cap R$ with $\mathrm{dist}(Q_{i_j}, S) > \sigma/2$. Since $\mathrm{dist}(Q_{i_j}, S) > \sigma/2$ as $i_j \to \infty$, let us assume Q_{i_j} converges to a limit $Q = (x', y') \in R - S$, implying $F(x', y'|\mathbf{P}) \neq 0$. Also let $Q_{i_j} = (x_{i_j}, y_{i_j})$. Then $F(x_{i_j}, y_{i_j}|\mathbf{P}_{i_j}) = 0$. Since $\mathbf{P}_{i_j} \to \mathbf{P}$ and $Q_{i_j} \to Q = (x', y')$, $F(x', y'|\mathbf{P}) = 0$, which contradicts the fact

that $F(x', y'|\mathbf{P}) \neq 0$. Therefore,

$$\sup_{Q \in S_i \cap R} \operatorname{dist}(Q, S) \to 0$$

So (A.4) and (A.4) imply

$$\operatorname{dist}_{\mathrm{H}}(S, S_i; R) = \max\Big\{ \sup_{P \in S \cap R} \operatorname{dist}(P, S_i), \sup_{Q \in S_i \cap R} \operatorname{dist}(Q, S) \Big\} \to 0$$

By combining (i) and (ii), we have

$$\operatorname{dist}_{\mathrm{H}}(S_i, S; R) \to 0 \quad \text{as} \quad i \to \infty.$$

for any finite window R. Now the proof of theorem is complete and we see that

$$S_i \to S \quad \text{as} \quad \mathbf{P}_i \to \mathbf{P}$$

The proof is complete. □

THEOREM A.7. *(**Divergence to Infinity**)If a sequence of parameter vectors converges to one of two poles* $(0, 0, 0, 0, 0, \pm 1)$ *or a point in the domain of imaginary parallel lines, the objective function representing sum of squares of orthogonal distances diverges to infinity.*

PROOF. (I)**Poles**

For a given set of n points, let d_{max} be the furthest distance between a point in the set and the origin. Pick a sequence of parameter vectors $P_k = (A_k, B_k, C_k, D_k, E_k, F_k)$ converging to $(0, 0, 0, 0, 0, 1)$ (the proof of $(0, 0, 0, 0, 0, -1)$ easily follows). Suppose $\varepsilon \ll \min(1/(M + d_{max}), 1/(M + d_{max})^2)$ where $M > 0$. For any point (x, y) with $dist((x, y), (0, 0)) < M + d_{max}$, since $A_k, B_k, C_k, D_k, E_k < \varepsilon$ for sufficiently large k, $|A_k x^2 + 2B_k xy + C_k y^2 + D_k x + E_k y| \ll 1$. Note that $F_k \approx 1$. So $A_k x^2 + 2B_k xy + C_k y^2 + D_k x + E_k y + F_k \neq 0$. It follows that any points (x, y) with $dist((x, y), (0, 0)) < M + d_{max}$ can not satisfy the equation. Thus the conic corresponding to $A_k x^2 + 2B_k xy + C_k y^2 + D_k x + E_k y + F_k = 0$ has a sum of squares of distances is larger than nM^2. Therefore, for any $nM^2 > 0$, if k is large enough, the sum of squares

of distances is greater $nM^2 >$. The objective function than diverges to infinity as $k \to \infty$.

(II)**Parallel lines**

The imaginary parallel lines is another type of empty solution for our model. Every sequence parameter vectors converging to a points in "IPL" corresponds to a sequence of conics for which the objective function diverges to infinity. The proof is similar to that of "two pole". The imaginary parallel lines are naturally described by the equation in a form of

$$(ax + by + c)^2 = d$$

where $d < 0$ (see detail in next section). Suppose a sequence of parameter vectors P_i converging to a limit P_0 corresponds to

$$(a_0x + b_0y + c_0)^2 = d_0 \qquad d_0 < 0$$

For sufficiently large i, the conic corresponding to P_i can be represented by an equation

$$a_1x^2 + b_1xy + c_1y^2 + d_1x + e_1y + f_1 + (a_0x + b_0y + c_0)^2 = d_0$$

where $a_1 \cdots f_1$ are extremely small. So for any point (x, y) close to a given set of n points,

$$a_1x^2 + b_1xy + c_1y^2 + d_1x + e_1y + f_1 \approx 0$$

Then (2) is approximately the same as (1) and (x, y) does not satisfy the equation (1). By the similar argument used for the proof of "two poles", the objective function for the sequence P_i diverges to infinity. □

The equation of imaginary parallel lines Here we study the quadratic equation for the imaginary parallel lines. Recall that a given equation

(A.94) $$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0,$$

corresponds to a pair of imaginary parallel lines when

(A.95)

$$Q = A^2 + B^2 + C^2 > 0 \qquad \Delta = \det \begin{bmatrix} A & B & D \\ B & C & E \\ D & E & F \end{bmatrix} = 0 \qquad J = \det \begin{bmatrix} A & B \\ B & C \end{bmatrix} = 0.$$

(A.96)
$$K = \det \begin{bmatrix} A & D \\ D & F \end{bmatrix} + \det \begin{bmatrix} C & E \\ E & F \end{bmatrix} > 0$$

(see section (4.2)).

We see that

$$\Delta = JF - E \cdot \det \begin{bmatrix} A & D \\ B & E \end{bmatrix} + D \cdot \det \begin{bmatrix} B & D \\ C & E \end{bmatrix} = -AE^2 + 2BDE - CD^2 = 0$$

Then

(A.97)
$$AE^2 + CD^2 = 2BDE.$$

Squaring both sides and using $J = 0$, we obtained

$$A^2E^4 + 2ACE^2D^2 + C^2D^4 = 4B^2D^2E^2 = 4ACD^2E^2$$

$$(AE^2 - CD^2)^2 = 0$$

. So $AE^2 = CD^2$. Since $Q = A^2 + B^2 + C^2 > 0$ and $B^2 = AC$, at least one of A and C is nonzero. Let us consider the following cases:

(I)$A \neq 0$ and $C = 0$.

It immediately follows that $B = 0$ and $E^2 = CD^2/A = 0$. Then (A.94) simplifies to

$$Ax^2 + 2Dx + F = 0.$$

which can be transformed into

(A.98)
$$(x + \frac{D}{A})^2 = -\frac{AF - D^2}{A^2}$$

where the right side $-\frac{AF-D^2}{A^2} = -\frac{K}{A^2} < 0$.

(II) $A, C \neq 0$.

Because of $AE^2 = CD^2$, we either have $D = E = 0$ or

$$A = tD^2 \qquad C = tE^2 \tag{A.99}$$

where $t \neq 0$. Furthermore, the sign of A and C must be the same. If the first case holds, (A.94) becomes

$$Ax^2 + 2Bxy + Cy^2 + F = Ax^2 \pm 2ACxy + Cy^2 + F = 0.$$

To make the positive coefficients for the quadratic term, we multiply both sides by $A + C$ and get

$$(\sqrt{(A+C)A}x \pm \sqrt{(A+C)C}y)^2 = -(A+C)F = -K < 0 \tag{A.100}$$

If the latter case holds, Substitute(A.99) into (A.97). Then

$$B = \frac{tD^2E^2 + tE^2D^2}{2DE} = tDE.$$

Rewrite (A.94) in terms of t, D, E and F as follows:

$$tD^2x^2 + 2tDExy + tE^2y^2 + 2Dx + 2Ey + F = 0$$

$$t(Dx + Ey)^2 + 2(Dx + Ey) + F = 0$$

. By making the left side a perfect square, we have

$$(Dx + Ey + \frac{1}{\sqrt{t}})^2 = \frac{1 - tF}{t^2}. \tag{A.101}$$

Also note that

$$K = AF - D^2 + CF - E^2 = tF(D^2 + E^2) - (D^2 + E^2) = (tF - 1)(D^2 + E^2) > 0.$$

and therefore $tF - 1 > 0$. So the right side of (A.101) is again negative.

In conclusion, we showed that every quadratic equation for a pair of imaginary parallel lines can be transformed into a standard form

$$(ax + by + c)^2 = d \quad d < 0.$$

which geometrically characterizes the imaginary parallel lines. We will use it for the proof in the proof of Theorem .

THEOREM A.8. *(**A basic fact about the sequence with a limit of coincident lines**) Suppose a sequence $\mathbf{P}_n$ of parameter vectors corresponding to real (not imaginary) conics, S_n, converges to a parameter vector $\mathbf{P} \in \mathbb{D}_{\mathrm{CL}}$ corresponding to a pair of coincident lines; the latter make a line in $\mathbb{R}^2$ which we denote by L. Then S_n gets closer and closer to L, as n grows. More precisely, for any rectangle*

$$R = \{-A \le x \le A, \quad -B \le y \le B\} \qquad (R \cap S_n, L \neq \emptyset)$$

($R \cap S_n$ should be nonempty for the maximum to be well defined. The theorem doesn't cover the example that a sequence of conics go off to infinity when $\mathbf{P}_n \to \mathbf{A}$.) we have

$$\max_{P \in S_n \cap R} \operatorname{dist}(P, L) \to 0 \qquad \text{as} \quad n \to \infty.$$

PROOF. By way of contradiction, let us assume that

$$\lim_{n \to \infty} \max_{P \in S_n \cap R} \operatorname{dist}(P, L) \neq 0.$$

or the limit does not even exists for some R.

Then there exists a $\varepsilon > 0$ such that

$$\forall N > 0 \qquad \max_{P \in S_n \cap R} \operatorname{dist}(P, L) > \varepsilon$$

for some $n > N$. So we can find a sequence of conics S_{n_m} such that

$$\max_{P \in S_{n_m} \cap R} \operatorname{dist}(P, L) = \operatorname{dist}(P_{n_m}, L) > \varepsilon$$

where $P_{n_m} \in S_{n_m} \cap R$. Since R is compact, let us assume that P_{n_m} converges to P_0. Clearly,

$$\lim_{n_m \to \infty} \max_{P \in S_{n_m} \cap R} \operatorname{dist}(P, L) = \lim_{n_m \to \infty} \operatorname{dist}(P_{n_m}, L) = \operatorname{dist}(P_0, L) > \varepsilon$$

Thus $P_0 \notin L$. Let $P_{n_m} = (x_{n_m}, y_{n_m})$. Note that the quadratic equation

$$F(x_{n_m}, y_{n_m} | \mathbf{P}_{n_m}) = 0$$

for all n_ms. So

$$\lim_{n_m\to\infty} F(x_{n_m}, y_{n_m}|\mathbf{P}_{n_m}) = F(x_0, y_0|\mathbf{P}) = 0$$

implying $P_0 \in L$. But remember $P_0 \notin L$ by our assumption. Therefore, we must have

$$\lim_{n\to\infty} \max_{P\in S_n\cap R} \quad \text{dist}(P, L) = 0.$$

□

A.5. Differentiability of the objective function on the sphere

THEOREM A.9. (5.4)*(**Differentiability of projection coordinates**) Let S be a conic and P a given point. Suppose (i) the point Q on the conic S closest to the given point P is unique and (ii) P is not the center of curvature of the conic S at the point Q. Then the coordinates x and y of the point Q are differentiable with respect to the conic's parameters.*

PROOF. Let $P = (x_0, y_0)$ be the given (fixed) point. The conic S is defined by a quadratic equation,

$$F(x, y; \mathbf{P}) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

where $\mathbf{P} = (A, B, C, D, E, F)$ represents the parameter vector. Let $Q = (u, v)$ denote he projection of P onto S. When the parameters A, B, C, D, E, F vary, the conic S changes, and so does the projection point Q. Thus its coordinates u and v become functions of the conic's parameters A, B, C, D, E, F.

By orthogonality of the line PQ to the conic S we have

$$F_u(u, v; \mathbf{P})(v - y_0) - F_v(u, v; \mathbf{P})(u - x_0) = 0, \tag{A.102}$$

where $\mathcal{P}_u$ and $\mathcal{P}_v$ denote partial derivatives of $\mathcal{P}$ with respect to u snd v. Also, since Q lies on the conic S, we have

$$F(u, v; \mathbf{P}) = 0. \tag{A.103}$$

The coordinates u and v are now specified, implicitly, by (A.102) and (A.103).

Next we apply Implicit Function Theorem. It guarantees that the coordinates u and v have continuous derivatives with respect to the parameters A, B, C, D, E, F provided a certain regularity condition is met. The regularity condition requires that the matrix of partial derivatives of the right hand sides of (A.102) and (A.103), with respect to u and v, is not singular. That is,

(A.104)
$$\det \begin{bmatrix} F_{uu}(v-y_0) - F_{uv}(u-x_0) - F_v & F_{uv}(v-y_0) - F_{vv}(u-x_0) + F_u \\ F_u & F_v \end{bmatrix} \neq 0.$$

Suppose that the determinant is zero. Then

$$(F_u F_{vv} - F_{uv}F_v)(u-x_0) + (F_v F_{uu} - F_{uv}F_u)(v-y_0) = F_u^2 + F_v^2 \tag{A.105}$$

According to the orthogonality relation (A.102), there exists a scalar t such that

$$(u-x_0) = tF_u \qquad (v-y_0) = tF_v \tag{A.106}$$

Substitution of (A.106) into (A.105) and solving for t give

$$t = \frac{F_u^2 + F_v^2}{F_u^2 F_{vv} - 2F_u F_v F_{uv} + F_v^2 F_{uu}}.$$

Note that

(A.107)
$$\text{dist}(Q,P) = \sqrt{(u-x_0)^2 + (v-y_0)^2} = |t|\sqrt{F_u^2 + F_v^2} = \frac{(F_u^2 + F_v^2)^{3/2}}{|F_u^2 F_{vv} - 2F_u F_v F_{uv} + F_v^2 F_{uu}|}.$$

This expression coincides with that for the radius of curvature of S at the point $Q = (u, v)$. So (A.104) can only hold if the given point $P = (x_0, y_0)$ happens to be exactly at the center of curvature of S. In all the other cases Implicit Function Theorem guarantees the existence of continuous derivatives of the coordinates u and v with respect to the conic's parameters.

Our theorem is now proved.

Lastly we give an explicit formula for the derivatives of u and v with respect to any parameter θ:

(A.108)

$$\begin{bmatrix} du/d\theta \\ dv/d\theta \end{bmatrix} = \begin{bmatrix} F_{uu}(v-y_0) - F_{uv}(u-x_0) - F_v & F_{uv}(v-y_0) - F_{vv}(u-x_0) + F_u \\ F_u & F_v \end{bmatrix}^{-1} \cdot \begin{bmatrix} F_{v\theta}(u-x_0) - F_{u\theta}(v-y_0) \\ -F_\theta \end{bmatrix}.$$

Here θ denotes an arbitrary component of the parameter vector $\mathbf{P}$, i.e., one can replace θ with A, B, etc. □

THEOREM A.10. *(**smoothness at centers of osculating circles**) Let S be a conic and P a given point. Suppose (i) the point Q on the conic S closest to the given point P is unique and (ii) P coincides with the center of curvature of the conic S at the point Q. Then the distance* $\operatorname{dist}(P, S)$ *is differentiable with respect to the conic's parameters.*

PROOF. . When the parameters of the conic S vary, the latter gets perturbed, and we denote the new (perturbed) conic by S' and the projection of the fixed point P onto the new conic by Q'. Clearly, Q' changes continuously with the parameters of the conic. Even if two distinct projections of P onto S' arise, both are close to the original projection Q (see an illustration below). Also, the derivative of the distance $d = \operatorname{dist}(P, S)$ with respect to the conic's parameters is given by the following general formula:

$$\nabla_{\mathbf{P}} d = \frac{\nabla_{\mathbf{P}} F(\bar{x}_i, \bar{y}_i; \mathbf{P})}{\|\nabla F(\bar{x}, \bar{y}; \mathbf{P})\|} \tag{A.109}$$

where $(\bar{x}, \bar{y})$ denote the coordinates of the projection point Q and

$$F(x, y; \mathbf{P}) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F$$

denotes the quadratic polynomial corresponding to the parameter vector $\mathbf{P} = (A, B, C, D, E, F)$. The formula (A.109) is derived in [**16**]. The numerator of the fraction in

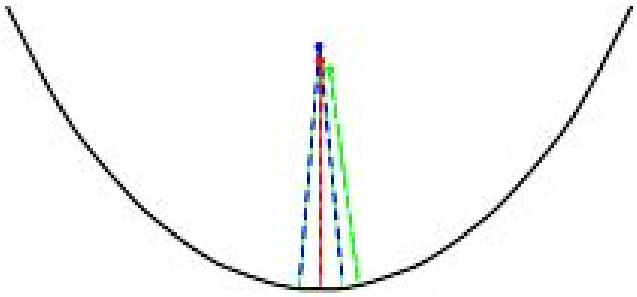


Figure A.10: Red point lies at the center of curvature of the conic. Blue point has two projections on the ellipse, but both are close to the projection of the red point

(A.109) contains the gradient of $\mathcal{P}$ with respect to the components of $\mathbf{P}$ and the denominator – the gradient of $\mathcal{P}$ with respect to x and y. Both gradients are taken at the projection point $Q = (\bar{x}, \bar{y})$. Clearly, all the elements of the fraction in (A.109) change continuously with Q (and hence with the conic's parameters). This proves that d is a smooth function of $\mathbf{P}$, i.e., it has continuous first order derivatives with respect to $\mathbf{P}$.

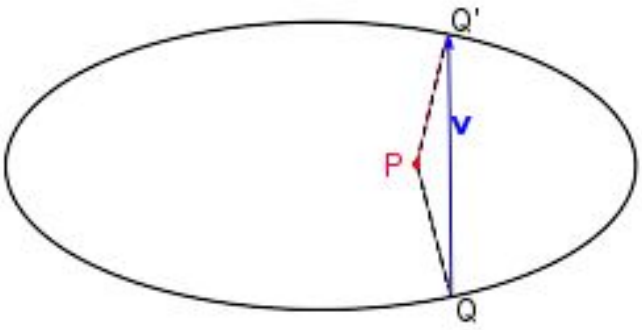
If the conic S is a circle and the data point P is at its center, it has multiple projections onto S, and this is precluded by the assumption (i) of the theorem. □

THEOREM A.11. (5.6)*(**Smoothness at local minima**) The objective function $\mathcal{F}$ is smooth at all its local minima. More precisely, the first order derivatives of $\mathcal{F}$, as well as those of the distances* $\mathrm{dist}(P_i, S)$*, exist and are continuous at all local minima.*

PROOF. Recall that singularities of $\mathcal{F}(S)$ are caused by two factors: (i) a data point $P_i = (x_i, y_i)$ happens to be at the center of curvature of the conic S and (ii) a data point $P_i = (x_i, y_i)$ has multiple projections on the conic, i.e., there are (at least) two distinct points $Q_i^{'}, Q_i^{''} \in S$ such that

$$\mathrm{dist}(P_i, S) = \mathrm{dist}(P_i, Q_i^{'}) = \mathrm{dist}(P_i, Q_i^{''}).$$

Now we deal with case (ii). Suppose a data point $P_i = (x_i, y_i)$ has two distinct projections, $Q_i^{'}$ and $Q_i^{''}$, onto the conic S. Recall that a translation of a conic S



along any vector $\mathbf{v}$ will be another conic. Let $\mathbf{v} = Q_i'Q_i''$, i.e., consider a vector from Q_i' to Q_i''. Then the translation of S in the direction of $\mathbf{v}$ will bring Q_i' closer to P_i and will move Q_i'' farther away from P_i. The translation of S in the direction of $-\mathbf{v}$ will have the opposite effect. In both cases the distance d_i from the data point P_i to the conic will decrease. The translation of the conic along the vector $\mathbf{v}$ corresponds to a directional derivative of the distance d_i in the parameter space. The above argument shows that this directional derivative is discontinuous: the distance d_i linearly decreases in both directions, whether the conic is shifted along $\mathbf{v}$ or along $-\mathbf{v}$. This creates a "peak" (local maximum) in the graph of the function d_i, and hence a "peak" in the graph of the objective function $\mathcal{F}$. The proof is completed. □

A.6. Objective function near boundaries

THEOREM A.12. *5.7 For any set of data points $P_1, \dots, P_n$ the global minimum of the objective function $\mathcal{F}$ belongs to the union*

$$\mathfrak{D}_{\mathcal{F},\mathrm{ESS}} = \mathbb{D}_{\mathrm{E}} \cup \mathbb{D}_{\mathrm{H}} \cup \mathbb{D}_{\mathrm{P}} \cup \mathbb{D}_{\mathrm{IL}} \cup \mathbb{D}_{\mathrm{PL}}. \tag{A.110}$$

If the objective function $\mathcal{F}$ has multiple global minima, then at least one of them belongs to the above union. This union cannot be shortened, i.e., for any conic S in this union of domains there exists a data set for which S provides the unique best fit.

PROOF. The existence of the best fitting conic has been shown in section (2.8). If that conic does not belong to the essential domain $\mathfrak{D}_{\mathcal{F},\mathrm{ESS}}$, then it belongs to one of

the remaining types of conics: $\mathbb{D}_{\mathrm{SL}}$ (single lines), $\mathbb{D}_{\mathrm{CL}}$ (coincident lines) or $\mathbb{D}_{\mathrm{SP}}$ (single points). Each of these conics is a proper subset of another conic from the essential domain $\mathfrak{D}_{\mathcal{F},\mathrm{ESS}}$. More precisely, intersecting lines or parallel lines suffice. So by the redundancy principle (see section (2.4)) there will be a best fitting conic from the union...

Now we prove that the union (A.110) cannot be shortened, i.e., no conic from the essential domain $\mathfrak{D}_{\mathcal{F},\mathrm{ESS}}$ can be discarded. If a conic S belongs to $\mathbb{D}_{\mathrm{E}} \cup \mathbb{D}_{\mathrm{H}} \cup \mathbb{D}_{\mathrm{P}}$, then it is a non-degenerate conic. Let $P_1, \ldots, P_5$ be any five distinct points on S. Then S interpolates those points, hence $\mathcal{F}(S) = 0$. No other conic can interpolate those five points (see wiki). Thus the best fitting conic S is unique.

Lastly, let a conic S belong to $\mathbb{D}_{\mathrm{IL}} \cup \mathbb{D}_{\mathrm{PL}}$. Then S is a pair of (intersecting or parallel) lines, say L_1 and L_2. Let P_1, P_2, P_3 be any three distinct points on L_1 and P_4, P_5 be any two distinct points on L_2 (if L_1 and L_2 intersect, we avoid the point of intersection when selecting $P_1, \ldots, P_5$, so these five points are truly distinct). Then S interpolates those five points, hence $\mathcal{F}(S) = 0$. They are not in general linear position, so they cannot be interpolated by a non-degenerate conic. Suppose our five points are interpolated by a degenerate conic S', i.e., by two other lines L_1' and L_2'. By the pigeonhole principle, one of them must contain at least three of our points, which is only possible if it contains P_1, P_2, P_3, hence it coincides with L_1. Now the other line must contain P_4 and P_5, so it coincides with L_2. □

A.7. Infinite moment of geometric parameters (General Case)

In this section we will derive the exact form of hessian matrix (A.129) used in the proof of infinite moment of geometric parameters in the section 6.4. Recall that the geometric parameters are center coordinates (C_x, C_y), two semi axis a, b and the angle θ between major axis and x axis. Our configuration consists of n points: three points in small squares B_i $(i = 1, 2, 3)$ of size $h^2 * h^2$ centered at $(2, 1)$, $(-2, 1)$, $(1, 0)$, one point in the small rectangle B_4 of size $h^2 * h$ centered at $(0, -1/3)$ and last $n-4$ points

in the square B_5 centered at $(-1, 0)$. All n points are assumed to be fixed except the fourth point in B_4. As it moves down from the top of the B_4 to the bottom ,the best fitting conic changes from ellipse to parabola, then become a hyperbola. The configuration is completely the same as the one in the proof of infinite moment for five points in the section 6.3, except that $n - 5$ points are added to the square B_5. Moreover, we introduced a new set of parameters: C_x, $1/(a + C_y)$, $a - C_y$, b^2/a and θ (denoted by $p_1, \ldots, p_5$ for which we will derive the hessian matrix.

Let (x, y) be the given point and (u, v) its projection on a given conic described by the equation $P(x, y; \theta)$ where θ is a parameter vector for the conic. Furthermore, let us denoted by d the distance between the given point and the conic. If α and β are two independent parameters for the conic, we have

(A.111)

$$d_\alpha d_\beta + d d_{\alpha\beta} = tP_{\alpha\beta} - (P_\alpha, tP_{u\alpha}, tP_{v\alpha}) \begin{pmatrix} 0 & P_u & P_v \\ P_u & 1 + tP_{uu} & tP_{uv} \\ P_v & tP_{uv} & 1 + tP_{vv} \end{pmatrix}^{-1} \begin{pmatrix} P_\beta \\ tP_{u\beta} \\ tP_{v\beta} \end{pmatrix}$$

where $t = \frac{d}{\sqrt{P_u^2 + P_v^2}}$. By simple geometry, the best fitting conic must pass through $B_1, \ldots, B_5$ and hence the distance d_i $(i = 1, \ldots, n)$ from each point to the curve are less than h^2. Furthermore, all derivatives P_{uu}, P_{uv}, P_{vv}, $P_{u\beta}$ etc are uniformly bounded by a constant M that may depend on n and h but not on point coordinates. Thus (A.112) can be reduced to

(A.112)

$$\mathcal{F}_{\alpha\beta}(\theta) = -\sum_1^n (P_\alpha, 0, 0) \begin{pmatrix} 0 & P_u & P_v \\ P_u & 1 & 0 \\ P_v & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} P_\beta \\ 0 \\ 0 \end{pmatrix} + \mathfrak{X}_h = \sum_1^n \frac{P_\alpha P_\beta}{P_u^2 + P_v^2} + \mathfrak{X}_h$$

where $\mathfrak{X}_h$ is a quantity that can be made arbitrarily small by decreasing h). So we only need the first derivatives of P with respect to each new parameters.

In section 6.3, we found the equation of the quadratic conic passing through $(1,0)$, $(-1,0)$, $(2,1)$, $(-2,1)$ and $(x_0, z-1/3)$ $(z \geqslant 0)$:

$$x^2 + ty^2 - (3+t)y - 1 = 0 \tag{A.113}$$

where t will be determined by the the point (x_0, y_0):

$$t = \frac{3z - x_0^2}{\frac{4}{9} - \frac{5}{3}z + z^2} \tag{A.114}$$

If $3z > x_0^2$, the quadratic curve corresponds to (A.113) is an ellipse with two semi axis

$$\hat{a} = \frac{\sqrt{9+10t+t^2}}{2t} \qquad \hat{b} = \frac{\sqrt{9+10t+t^2}}{2\sqrt{t}} \tag{A.115}$$

and y coordinates of two focuses (x coordinates are 0)

$$\frac{3+t}{2t} \pm \sqrt{\frac{9+10t+t^2}{4t^2} - \frac{9+10t+t^2}{4t}} \tag{A.116}$$

So

$$\frac{\hat{b}^2}{\hat{a}} = \frac{\sqrt{9+10u+u^2}}{2} \to \frac{3}{2} \quad \text{as} \quad u \to 0 \tag{A.117}$$

As $t \to 0$, the y coordinate of upper focus (k_1, h_1)

$$h_1 = \frac{3+t}{2t} + \sqrt{\frac{9+10t+t^2}{4t^2} - \frac{9+10t+t^2}{4t}} \to +\infty \tag{A.118}$$

and the y coordinate of lower focus (k_2, h_2)

$$h_2 = \frac{3+t}{2t} - \sqrt{\frac{9+10t+t^2}{4t^2} - \frac{9+10t+t^2}{4t}} \to \frac{12}{5} \tag{A.119}$$

The above results can be considered as a approximation to our construction for the general case. For convenience, we will use the following shorthand notations:

$$d_{f1} = \sqrt{(u-k_1)^2 + (v-h_1)^2} \to \infty$$

$$d_{f2} = \sqrt{(u-k_2)^2 + (v-h_2)^2} \approx \sqrt{u^2 + (v - \frac{12}{5})^2}$$

as ellipse changes to a parabola. Also we use following relations:

$$A = \frac{1}{2}(1/p_2 + p_3) \quad C_y = \frac{1}{2}(1/p_2 - p_3) \quad B = \sqrt{\frac{1}{2}(1/p_2 + p_3)p_4} \tag{A.120}$$

Remember that a ellipse is described by the equation:

$$\sqrt{(u-x_1)^2 + (v-y_1)^2} + \sqrt{(u-x_2)^2 + (v-y_2)^2} - 2a = 0 \tag{A.121}$$

By differentiating (A.121) with respect to each new parameters when the best fitting curve becomes a parabola, we have

$$P_{p_1} = -\frac{u-k_1}{d_{f1}} - \frac{u-k_2}{d_{f2}} \to -\frac{u-k_2}{d_{f2}} + \mathcal{X}_1 \tag{A.122}$$

$$P_{p_2} = -\frac{9}{16}\frac{v-h_2}{d_{f2}} - \frac{9}{16} + \frac{u^2}{2} + \mathcal{X}_2 \tag{A.123}$$

$$P_{p_3} = \frac{v-h_2}{d_{f2}} - 1 + \mathcal{X}_3 \tag{A.124}$$

$$P_{p_4} = \frac{1}{2}(-1 - \frac{v-h_2}{d_{f2}}) + \mathcal{X}_4 \tag{A.125}$$

$$P_{p_4} = u(1 + \frac{5}{12d_{f2}}) + \mathcal{X}_5 \tag{A.126}$$

where $\mathcal{X}_1, \ldots, \mathcal{X}_h$ are small quantities (that can be made arbitrarily small by decreasing h). Furthermore, we can also obtain

$$P_u = u(u^2 + (v - \frac{5}{12})^2)^{-1/2} \tag{A.127}$$

$$P_v = (u^2 + (v - \frac{5}{12})^2)^{-1/2}(v - \frac{5}{12}) - 1 \tag{A.128}$$

Now using (A.112) for each element of Hessian matrix, we obtain

(A.129)

$$H = \begin{pmatrix} \frac{232}{325} + \frac{8}{13}n & -\frac{20}{39} + \frac{4}{39}n & \frac{60}{13} - \frac{12}{13}n & \frac{40}{39} - \frac{8}{39}n & -\frac{236}{65} - \frac{12}{13}n \\ -\frac{20}{39} + \frac{4}{39}n & \frac{13162}{2925} + \frac{2}{117}n & -\frac{682}{325} - \frac{2}{13}n & -\frac{6356}{2925} - \frac{4}{117}n & \frac{10}{13} - \frac{2}{13}n \\ \frac{60}{13} - \frac{12}{13}n & -\frac{682}{325} - \frac{2}{13}n & -\frac{232}{325} + \frac{18}{13}n & \frac{116}{325} + \frac{4}{13}n & -\frac{90}{13} + \frac{18}{13}n \\ \frac{40}{39} - \frac{8}{39}n & -\frac{6356}{2925} - \frac{4}{117}n & \frac{116}{325} + \frac{4}{13}n & \frac{2728}{2925} + \frac{8}{117}n & \frac{-20}{13} + \frac{4}{13}n \\ -\frac{236}{65} - \frac{12}{13}n & \frac{10}{13} - \frac{2}{13}n & -\frac{90}{13} + \frac{18}{13}n & -\frac{20}{13} + \frac{4}{13}n & \frac{154}{13} + \frac{18}{13}n \end{pmatrix} + \chi$$

(the above result is also checked by Maple) where χ can be made as small as possible by decreasing h.