Ellipse Fitting with Hyperaccuracy

Kenichi Kanatani

Department of Computer Science,
Okayama University, Okayama 700-8530, Japan
kanatani@suri.it.okayama-u.ac.jp

Abstract. For fitting an ellipse to a point sequence, ML (maximum likelihood) has been regarded as having the highest accuracy. In this paper, we demonstrate the existence of a "hyperaccurate" method which outperforms ML. This is made possible by error analysis of ML followed by subtraction of high-order bias terms. Since ML nearly achieves the theoretical accuracy bound (the KCR lower bound), the resulting improvement is very small. Nevertheless, our analysis has theoretical significance, illuminating the relationship between ML and the KCR lower bound.

1 Introduction

Circular and spherical objects in the scene are generally projected onto ellipses on the image plane, and their 3-D shapes and positions can be computed from their images [9]. For this reason, fitting ellipses (including circles) to a point sequence is one of the first steps of various vision applications, and numerous papers have been written on this subject. They are classified into two categories:

1. How can we judge whether a sequence of edge points entirely consists of points on an ellipse or it contains other points ("outliers")?
2. How can we fit the equation of an ellipse to a sequence of points known to be on an ellipse as accurately as possible?

For the first task, many algorithms and their efficient implementation techniques have been tested. There exists an abundance of literature on the second task, too. Most of the proposed methods were based on heuristics combining voting and least squares in many different forms [3, 4, 15, 20, 21, 22], but there are also theoretical treatments, mainly by statisticians, regarding the problem as statistical estimation [1, 2, 5, 16, 17, 19, 23]. However, their major concern is the consistency and efficiency of the estimator in the asymptotic limit as the number of points increases.

A contrasting approach was presented by Kanatani [11], who generalized ellipse fitting into an abstract framework, which he called *geometric fitting*. Having actual image processing in mind, he pursued fitting schemes whose accuracy rapidly increases as the noise level decreases for a fixed number of points. He asserted that such methods can tolerate larger image processing uncertainty for a desired accuracy level [13].

A. Leonardis, H. Bischof, and A. Pinz (Eds.): ECCV 2006, Part I, LNCS 3951, pp. 484–495, 2006.

In his framework, a lower bound on the covariance matrix of the estimator is obtained [11, 12]. Chernov and Lesort [6] called it the *KCR* (*Kanatani-Cramer-Rao*) *lower bound* and showed that it can be derived under a weaker assumption.

It can be shown that *ML* (*maximum likelihood*) can attain that bound except for higher order terms in the noise level [6, 11, 13]. It has turned out that all existing iterative linear computing schemes, such as *renormalization*[1] [10, 11, 14], *HEIV* [18], and *FNS* [7], has accuracy equivalent to ML [13]. It has been experimentally confirmed that these methods indeed attain high accuracy very close to the KCR lower bound.

We say that an estimation method has *hyperaccuracy* if it outperforms ML. In this paper, we demonstrate that there *does* exist a hyperaccurate method. Since ML nearly achieves the KCR lower bound, the accuracy improvement is very small. Nevertheless, our analysis has theoretical significance, illuminating the relationship between ML and the KCR lower bound.

2 KCR Lower Bound for Ellipse Fitting

We want to fit an ellipse to N points $\{(x_\alpha, y_\alpha)\}$, $\alpha = 1, ..., N$. An ellipse is represented by

$$Ax^2 + 2Bxy + Cy^2 + 2f_0(Dx + Ey) + Ff_0^2 = 0, \tag{1}$$

where f_0 is an arbitrary scaling constant[2]. If we define

$$\boldsymbol{u} = \begin{pmatrix} A & B & C & D & E & F \end{pmatrix}^\top, \qquad \boldsymbol{\xi} = \begin{pmatrix} x^2 & 2xy & y^2 & 2f_0x & 2f_0y & f_0^2 \end{pmatrix}^\top, \tag{2}$$

eq. (1) is written as

$$(\boldsymbol{u}, \boldsymbol{\xi}) = 0. \tag{3}$$

Throughout this paper, we denote the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ by $(\boldsymbol{a}, \boldsymbol{b})$. Since the magnitude of the vector $\boldsymbol{u}$ is indeterminate, we adopt normalization $\|\boldsymbol{u}\| = 1$. Geometrically, eq. (3) describes a hyperplane in the 6-dimensional space $\mathcal{R}^6$ of the variable vector $\boldsymbol{\xi}$. The N points $\{(x_\alpha, y_\alpha)\}$, $\alpha = 1, ..., N$, can be regarded as points in $\mathcal{R}^6$ via the embedding $\boldsymbol{\xi} : \mathcal{R}^2 \to \mathcal{R}^6$ defined by the second of eqs. (2). Thus, ellipse fitting in $\mathcal{R}^2$ is converted to hyperplane fitting in $\mathcal{R}^6$.

Remark. Eq. (1) describes not necessarily an ellipse but also a parabola, a hyperbola, and their degeneracies (e.g., two lines), generically called a *conic*. For this reason, fitting a curve in the form of eq. (1) is often called *conic fitting* [9]. Even if the points $\{(x_\alpha, y_\alpha)\}$ are sampled from an ellipse, the fitted equation may define a hyperbola or other curves in the presence of large noise, and a technique for preventing this has been proposed [8]. Here, we do not impose any constraints to prevent non-ellipses, assuming that noise is sufficiently small.

---

[1] The program is available at `http://www.suri.it.okayama-u.ac.jp`

[2] One can set $f_0 = 1$ unless the data have too large magnitudes, in which case a large value of f_0 would stabilize numerical computation.

Suppose each point (x_α, y_α) is perturbed from its true position $(\bar{x}_\alpha, \bar{y}_\alpha)$ by Gaussian noise of mean 0 and standard deviation σ in each component independently. Then, the covariance matrix of $\boldsymbol{\xi}_\alpha$ has the form $4\sigma^2 V_0[\boldsymbol{\xi}_\alpha]$, where $V_0[\boldsymbol{\xi}_\alpha]$, which we call the *normalized covariance matrix*, is given, after omitting higher order terms[3] in σ, by

$$V_0[\boldsymbol{\xi}_\alpha] = \begin{pmatrix} \bar{x}_\alpha^2 & \bar{x}_\alpha\bar{y}_\alpha & 0 & f_0\bar{x}_\alpha & 0 & 0 \\ \bar{x}_\alpha\bar{y}_\alpha & \bar{x}_\alpha^2 + \bar{y}_\alpha^2 & \bar{x}_\alpha\bar{y}_\alpha & f_0\bar{y}_\alpha & f_0\bar{x}_\alpha & 0 \\ 0 & \bar{x}_\alpha\bar{y}_\alpha & \bar{y}_\alpha^2 & 0 & f_0\bar{y}_\alpha & 0 \\ f_0\bar{x}_\alpha & f_0\bar{y}_\alpha & 0 & f_0^2 & 0 & 0 \\ 0 & f_0\bar{x}_\alpha & f_0\bar{y}_\alpha & 0 & f_0^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{4}$$

Since $\boldsymbol{\xi}_\alpha$ has only 2 degrees of freedom (i.e., x_α and y_α), $V_0[\boldsymbol{\xi}_\alpha]$ has rank 2.

Let $\hat{\boldsymbol{u}}$ be an estimator of $\boldsymbol{u}$ obtained by some means. Its accuracy is measured by the following covariance matrix:

$$V[\hat{\boldsymbol{u}}] = E[(\boldsymbol{P}_{\mathbf{u}}\hat{\boldsymbol{u}})(\boldsymbol{P}_{\mathbf{u}}\hat{\boldsymbol{u}})^\top]. \tag{5}$$

Here, $E[\,\cdot\,]$ denotes expectation with respect to the noise in the data $\{(x_\alpha, y_\alpha)\}$, and $\boldsymbol{P}_{\mathbf{u}}$ is the projection matrix ($\boldsymbol{I}$ denotes the unit matrix)

$$\boldsymbol{P}_{\mathbf{u}} = \boldsymbol{I} - \boldsymbol{u}\boldsymbol{u}^\top, \tag{6}$$

which projects $\hat{\boldsymbol{u}}$ onto the hyperplane orthogonal to $\boldsymbol{u}$. Since the parameter vector $\boldsymbol{u}$ is normalized to unit norm, its domain is the unit sphere $\mathcal{S}^5$ in $\mathcal{R}^6$. Following the approach of Kanatani [11], we focus on the asymptotic limit of small noise and identify the domain of the errors with the tangent hyperplane to $\mathcal{S}^5$ at $\boldsymbol{u}$. Namely, we evaluate the error after projecting it onto that hyperplane. Thus, the covariance matrix $V[\hat{\boldsymbol{u}}]$ is a singular matrix of rank 5.

In this setting, Kanatani [11, 13] proved that if $\boldsymbol{\xi}_\alpha$ is regarded as an independent Gaussian random variable of mean $\bar{\boldsymbol{\xi}}_\alpha$ and covariance matrix $V[\boldsymbol{\xi}_\alpha]$, the following inequality holds for an arbitrary unbiased estimator $\hat{\boldsymbol{u}}$ of $\boldsymbol{u}$:

$$V[\hat{\boldsymbol{u}}] \succ \Big(\sum_{\alpha=1}^{N} \frac{\bar{\boldsymbol{\xi}}_\alpha\bar{\boldsymbol{\xi}}_\alpha^\top}{(\boldsymbol{u}, V[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}\Big)^-. \tag{7}$$

Here, $\succ$ means that the difference of the left-hand side from the right is positive semidefinite, and the superscript $-$ denotes the generalized inverse (of rank 5).

Chernov and Lesort [6] called the right-hand side of eq. (7) the *KCR (Kanatani-Cramer-Rao) lower bound* and showed that it holds except for terms of $O(\sigma^4)$ even if $\hat{\boldsymbol{u}}$ is not unbiased; it is sufficient that $\hat{\boldsymbol{u}}$ is "consistent" in the sense that $\hat{\boldsymbol{u}} \to \boldsymbol{u}$ as $\sigma \to 0$.

[3] We confirmed by experiment that inclusion of the omitted higher order terms has no noticeable effects in our numerical results shown later.

3 Maximum Likelihood Estimation

The best known method for solving the above problem is the *least squares* (or *algebraic distance minimization*), minimizing

$$J_{\rm LS} = \sum_{\alpha=1}^{N} (\boldsymbol{u}, \boldsymbol{\xi}_\alpha)^2. \tag{8}$$

This is a quadratic form $J_{\rm LS} = (\boldsymbol{u}, \boldsymbol{M}_{\rm LS}\boldsymbol{u})$ in $\boldsymbol{u}$ if we define

$$\boldsymbol{M}_{\rm LS} = \sum_{\alpha=1}^{N} \boldsymbol{\xi}_\alpha \boldsymbol{\xi}_\alpha^\top. \tag{9}$$

Hence, the solution $\hat{\boldsymbol{u}}_{\rm LS}$ is the unit eigenvector of $\boldsymbol{M}_{\rm LS}$ for the smallest eigenvalue. However, the solution $\boldsymbol{u}_{\rm LS}$ is known to have large statistical bias [11].

If $\boldsymbol{\xi}_\alpha$ is regarded as an independent Gaussian random variable of mean $\bar{\boldsymbol{\xi}}_\alpha$ and covariance matrix $V[\boldsymbol{\xi}_\alpha]$, *ML* (*maximum likelihood*) is to minimize the sum of the square Mahalanobis distances of the data points $\boldsymbol{\xi}_\alpha$ to the hyperplane to be fitted, minimizing

$$J = \sum_{\alpha=1}^{N} (\boldsymbol{\xi}_\alpha - \bar{\boldsymbol{\xi}}_\alpha, V_0[\boldsymbol{\xi}_\alpha]^- (\boldsymbol{\xi}_\alpha - \bar{\boldsymbol{\xi}}_\alpha)), \tag{10}$$

subject to the constraint $(\boldsymbol{u}, \bar{\boldsymbol{\xi}}_\alpha) = 0$, $\alpha = 1, ..., N$. We can use $V_0[\boldsymbol{\xi}_\alpha]$ instead of the full covariance matrix $4\sigma^2 V_0[\boldsymbol{\xi}_\alpha]$, because the solution is unchanged if $V_0[\boldsymbol{\xi}_\alpha]$ is multiplied by a positive constant. Introducing Lagrange multipliers for the constraint $(\boldsymbol{u}, \bar{\boldsymbol{\xi}}_\alpha) = 0$, we can reduce the problem to unconstrained minimization of the following function [7, 11, 18]:

$$J = \sum_{\alpha=1}^{N} \frac{(\boldsymbol{u}, \boldsymbol{\xi}_\alpha)^2}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}. \tag{11}$$

By differentiation with respect to $\boldsymbol{u}$, we have

$$\nabla_{\mathbf{u}} J = \sum_{\alpha=1}^{N} \frac{2(\boldsymbol{\xi}_\alpha, \boldsymbol{u})\boldsymbol{\xi}_\alpha}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})} - \sum_{\alpha=1}^{N} \frac{2(\boldsymbol{\xi}_\alpha, \boldsymbol{u})^2 V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u}}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}. \tag{12}$$

The ML estimator $\hat{\boldsymbol{u}}$ is obtained by solving $\nabla_{\mathbf{u}} J = \boldsymbol{0}$, or

$$\boldsymbol{M}\boldsymbol{u} = \boldsymbol{L}\boldsymbol{u}, \tag{13}$$

$$\boldsymbol{M} = \sum_{\alpha=1}^{N} \frac{\boldsymbol{\xi}_\alpha \boldsymbol{\xi}_\alpha^\top}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}, \qquad \boldsymbol{L} = \sum_{\alpha=1}^{N} \frac{(\boldsymbol{\xi}_\alpha, \boldsymbol{u})^2 V_0[\boldsymbol{\xi}_\alpha]}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}. \tag{14}$$

The FNS of Chojnacki et al. [7] solves eq. (13) by iteratively computing eigenvalue problems; the HEIV of Leedan and Meer [18] iteratively computes generalized eigenvalue problems. In theory, the renormalization of Kanatani [11] also solves eq. (14) with the same accuracy as the FNS and the HEIV [13].

4 Error Analysis of ML

Substituting $\boldsymbol{\xi}_\alpha = \bar{\boldsymbol{\xi}}_\alpha + \Delta\boldsymbol{\xi}_\alpha$ in the matrix $\boldsymbol{M}$ in eqs. (14), we obtain

$$\boldsymbol{M} = \bar{\boldsymbol{M}} + \Delta_1\boldsymbol{M} + \Delta_2\boldsymbol{M}, \tag{15}$$

$$\Delta_1\boldsymbol{M} = \sum_{\alpha=1}^{N} \frac{\Delta\boldsymbol{\xi}_\alpha\bar{\boldsymbol{\xi}}_\alpha^\top + \bar{\boldsymbol{\xi}}_\alpha\Delta\boldsymbol{\xi}_\alpha^\top}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}, \qquad \Delta_2\boldsymbol{M} = \sum_{\alpha=1}^{N} \frac{\Delta\boldsymbol{\xi}_\alpha\Delta\boldsymbol{\xi}_\alpha^\top}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}, \tag{16}$$

where $\bar{\boldsymbol{M}}$ is the value of the matrix $\boldsymbol{M}$ defined by the true values $\{\bar{\boldsymbol{\xi}}_\alpha\}$ of $\{\boldsymbol{\xi}_\alpha\}$. The matrix $\boldsymbol{L}$ in eqs. (14) is written as

$$\boldsymbol{L} = \sum_{\alpha=1}^{N} \frac{(\bar{\boldsymbol{\xi}}_\alpha + \Delta\boldsymbol{\xi}_\alpha, \boldsymbol{u})^2 V_0[\boldsymbol{\xi}_\alpha]}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2} = \sum_{\alpha=1}^{N} \frac{(\Delta\boldsymbol{\xi}_\alpha, \boldsymbol{u})^2 V_0[\boldsymbol{\xi}_\alpha]}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2} = \Delta_2\boldsymbol{L}. \tag{17}$$

Letting $\boldsymbol{u}$ be the noise-free value of the solution, we expand the ML estimator $\hat{\boldsymbol{u}}$ in the form

$$\hat{\boldsymbol{u}} = \boldsymbol{u} + \Delta_1\boldsymbol{u} + \Delta_2\boldsymbol{u} + \cdots, \tag{18}$$

where $\Delta_k\boldsymbol{u}$ denotes terms which contain kth powers of the components of $\Delta\boldsymbol{\xi}_\alpha$ having a magnitude of $O(\sigma^k)$. Substituting eq. (18) into eq. (13), we obtain

$$\begin{aligned}(\bar{\boldsymbol{M}} + \Delta_1\boldsymbol{M} + \Delta_1^*\boldsymbol{M} + \Delta_2\boldsymbol{M} + \Delta_2^*\boldsymbol{M} + \cdots)(\boldsymbol{u} + \Delta_1\boldsymbol{u} + \Delta_2\boldsymbol{u} + \cdots) \\ = \Delta_2\boldsymbol{L}(\boldsymbol{u} + \Delta_1\boldsymbol{u} + \Delta_2\boldsymbol{u} + \cdots),\end{aligned} \tag{19}$$

where $\Delta_1^*\boldsymbol{M}$ and $\Delta_2^*\boldsymbol{M}$ are, respectively, the perturbation terms arising by replacing $\boldsymbol{u}$ in the denominator $(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})$ in $\bar{\boldsymbol{M}}$ and $\Delta_1\boldsymbol{M}$ by $\hat{\boldsymbol{u}}$ (the corresponding perturbation of $\Delta_2\boldsymbol{M}$ is of $O(\sigma^3)$). They have the form

$$\Delta_1^*\boldsymbol{M} = -2\sum_{\alpha=1}^{N} \frac{((\Delta_1\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u}) + O(\sigma^2))\bar{\boldsymbol{\xi}}_\alpha\bar{\boldsymbol{\xi}}_\alpha^\top}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}, \tag{20}$$

$$\Delta_2^*\boldsymbol{M} = -2\sum_{\alpha=1}^{N} \frac{((\Delta_1\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u}) + O(\sigma^2))(\Delta\boldsymbol{\xi}_\alpha\bar{\boldsymbol{\xi}}_\alpha^\top + \bar{\boldsymbol{\xi}}_\alpha\Delta\boldsymbol{\xi}_\alpha^\top)}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}. \tag{21}$$

Equating terms of $O(1)$, $O(\sigma)$, and $O(\sigma^2)$ on both sides of eq. (19), we obtain the following expressions (we omit the derivation):

$$\Delta_1\boldsymbol{u} = -\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\boldsymbol{u} \tag{22}$$

$$\begin{aligned}\Delta_2\boldsymbol{u} = &-\bar{\boldsymbol{M}}^-\Delta_2\boldsymbol{M}\boldsymbol{u} + \bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\boldsymbol{u} + \bar{\boldsymbol{M}}^-\Delta_1^*\boldsymbol{M}\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\boldsymbol{u} \\ &-\bar{\boldsymbol{M}}^-\Delta_2^*\boldsymbol{M}\boldsymbol{u} + \bar{\boldsymbol{M}}^-\Delta_2\boldsymbol{L}\boldsymbol{u} - \|\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\boldsymbol{u}\|^2\boldsymbol{u}.\end{aligned} \tag{23}$$

From the first of eqs. (14), we have $\bar{\boldsymbol{M}}\boldsymbol{u} = \boldsymbol{0}$ and hence $\bar{\boldsymbol{M}}^-\boldsymbol{u} = \boldsymbol{0}$. It follows that terms on the right-hand sides of eqs. (22) and (23) are orthogonal to $\boldsymbol{u}$ except the last term $-\|\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\boldsymbol{u}\|^2\boldsymbol{u}$, which is parallel to $\boldsymbol{u}$, accounting for the normalization $\|\boldsymbol{u}\| = 1$ (Fig. 1).

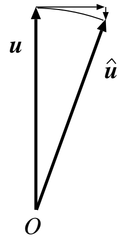


Fig. 1. The orthogonal and the parallel components of the error in $\hat{\boldsymbol{u}}$

It can be seen that the first order error $\Delta\boldsymbol{u}_1$ yields variations corresponding to the KCR lower bound. In fact, we have

$$\begin{aligned} E[\Delta_1\boldsymbol{u}\Delta_1\boldsymbol{u}^\top] &= E[\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\boldsymbol{u}\boldsymbol{u}^\top\Delta_1\boldsymbol{M}\bar{\boldsymbol{M}}^-] \\ &= E[\bar{\boldsymbol{M}}^-\sum_{\alpha=1}^N \frac{\Delta\boldsymbol{\xi}_\alpha\bar{\boldsymbol{\xi}}_\alpha^\top + \bar{\boldsymbol{\xi}}_\alpha\Delta\boldsymbol{\xi}_\alpha^\top}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}\boldsymbol{u}\boldsymbol{u}^\top\sum_{\beta=1}^N \frac{\Delta\boldsymbol{\xi}_\beta\bar{\boldsymbol{\xi}}_\beta^\top + \bar{\boldsymbol{\xi}}_\beta\Delta\boldsymbol{\xi}_\beta^\top}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\beta]\boldsymbol{u})}\bar{\boldsymbol{M}}^-] \\ &= \bar{\boldsymbol{M}}^-\sum_{\alpha,\beta=1}^N \frac{(\boldsymbol{u}, E[\Delta\boldsymbol{\xi}_\alpha\Delta\boldsymbol{\xi}_\beta^\top]\boldsymbol{u})\bar{\boldsymbol{\xi}}_\alpha\bar{\boldsymbol{\xi}}_\beta^\top}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\beta]\boldsymbol{u})}\bar{\boldsymbol{M}}^- \\ &= \bar{\boldsymbol{M}}^-\sum_{\alpha=1}^N \frac{4\sigma^2\bar{\boldsymbol{\xi}}_\alpha\bar{\boldsymbol{\xi}}_\alpha^\top}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}\bar{\boldsymbol{M}}^- = 4\sigma^2\bar{\boldsymbol{M}}^-\bar{\boldsymbol{M}}\bar{\boldsymbol{M}}^- = 4\sigma^2\bar{\boldsymbol{M}}^-, \end{aligned} \tag{24}$$

where we have used the identity[4] $E[\Delta\boldsymbol{\xi}_\alpha\Delta\boldsymbol{\xi}_\beta^\top] = 4\sigma^2\delta_{\alpha\beta}V_0[\boldsymbol{\xi}_\alpha]$, which is a consequence of our assumption that the noise in each $\boldsymbol{x}_\alpha$ is independent.

From the definition of $\bar{\boldsymbol{M}}$ and $V_0[\boldsymbol{\xi}_\alpha]$, we can see that eq. (24) coincides with the KCR lower bound. Adding the second order error $\Delta_2\boldsymbol{u}$ affects this only by $O(\sigma^4)$, since expectation of odd powers of $\Delta\boldsymbol{\xi}_\alpha$ is 0 due to the symmetry of the noise distribution. Thus, as pointed out by Kanatani [11] and Chernov and Lesort [6], the covariance matrix of ML attains the KCR lower bound except for $O(\sigma^4)$. We now examine the effect of the second order error $\Delta_2\boldsymbol{u}$.

5 Bias Evaluation for ML

Since $E[\Delta\boldsymbol{\xi}_\alpha] = \boldsymbol{0}$, we have $E[\Delta_1\boldsymbol{M}] = \boldsymbol{O}$. Hence, the first order error $\Delta_1\boldsymbol{u}$ is "unbiased". So, we evaluate the bias of the second order error $\Delta_2\boldsymbol{u}$. The expectation of $\Delta_2\boldsymbol{M}$ is

$$E[\Delta_2\boldsymbol{M}] = \sum_{\alpha=1}^N \frac{E[\Delta\boldsymbol{\xi}_\alpha\Delta\boldsymbol{\xi}_\alpha^\top]}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})} = \sum_{\alpha=1}^N \frac{4\sigma^2 V_0[\boldsymbol{\xi}_\alpha]}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})} = 4\sigma^2\boldsymbol{N}, \tag{25}$$

where we define

$$\boldsymbol{N} = \sum_{\alpha=1}^N \frac{V_0[\boldsymbol{\xi}_\alpha]}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}. \tag{26}$$

[4] The symbol $\delta_{\alpha\beta}$ is the Kronecker delta, taking on 1 for $\alpha = \beta$ and 0 otherwise.

The expectation of $\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u}$ is

$$\begin{aligned}
&E[\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u}] \\
&= E[\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\frac{\bar{\boldsymbol{\xi}}_\alpha\Delta\boldsymbol{\xi}_\alpha^\top+\Delta\boldsymbol{\xi}_\alpha\bar{\boldsymbol{\xi}}_\alpha^\top}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}\bar{\boldsymbol{M}}^{-}\sum_{\beta=1}^{N}\frac{\bar{\boldsymbol{\xi}}_\beta\Delta\boldsymbol{\xi}_\beta^\top+\Delta\boldsymbol{\xi}_\beta\bar{\boldsymbol{\xi}}_\beta^\top}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\beta]\boldsymbol{u})}\boldsymbol{u}] \\
&= \sum_{\alpha,\beta=1}^{N}\frac{\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha(\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\beta)^\top E[\Delta\boldsymbol{\xi}_\alpha\Delta\boldsymbol{\xi}_\beta^\top]\boldsymbol{u}+\bar{\boldsymbol{M}}^{-}E[\Delta\boldsymbol{\xi}_\alpha\Delta\boldsymbol{\xi}_\beta^\top]\boldsymbol{u}(\bar{\boldsymbol{\xi}}_\alpha,\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\beta)}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})(\boldsymbol{u},V_0[\boldsymbol{\xi}_\beta]\boldsymbol{u})} \\
&= 4\sigma^2\sum_{\alpha=1}^{N}\frac{(\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha,V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha+(\bar{\boldsymbol{\xi}}_\alpha,\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha)\bar{\boldsymbol{M}}^{-}V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u}}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}. \qquad (27)
\end{aligned}$$

The expectation of $\Delta_1^*\boldsymbol{M}\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u}$ is

$$\begin{aligned}
E[\Delta_1^*\boldsymbol{M}\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u}] &= -2\sum_{\alpha=1}^{N}\frac{E[(\Delta_1\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})\bar{\boldsymbol{\xi}}_\alpha(\bar{\boldsymbol{\xi}}_\alpha,\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u})]}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2} \\
&= -2\sum_{\alpha=1}^{N}\frac{E[(\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})(\bar{\boldsymbol{\xi}}_\alpha,\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u})\bar{\boldsymbol{\xi}}_\alpha]}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2} \\
&= -2\sum_{\alpha=1}^{N}\frac{(\bar{\boldsymbol{M}}^{-}V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u},E[(\Delta_1\boldsymbol{M}\boldsymbol{u})(\Delta_1\boldsymbol{M}\boldsymbol{u})^\top]\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha)\bar{\boldsymbol{\xi}}_\alpha}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}. \qquad (28)
\end{aligned}$$

We can evaluate $E[(\Delta_1\boldsymbol{M}\boldsymbol{u})(\Delta_1\boldsymbol{M}\boldsymbol{u})^\top]$ as follows:

$$\begin{aligned}
E[(\Delta_1\boldsymbol{M}\boldsymbol{u})(\Delta_1\boldsymbol{M}\boldsymbol{u})^\top] &= E[\sum_{\alpha=1}^{N}\frac{\bar{\boldsymbol{\xi}}_\alpha(\Delta\boldsymbol{\xi}_\alpha,\boldsymbol{u})}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}\sum_{\beta=1}^{N}\frac{\bar{\boldsymbol{\xi}}_\beta^\top(\Delta\boldsymbol{\xi}_\beta,\boldsymbol{u})}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\beta]\boldsymbol{u})}] \\
= \sum_{\alpha,\beta=1}^{N}\frac{(\boldsymbol{u},E[\Delta\boldsymbol{\xi}_\alpha\Delta\boldsymbol{\xi}_\beta^\top],\boldsymbol{u})\bar{\boldsymbol{\xi}}_\alpha\bar{\boldsymbol{\xi}}_\beta^\top}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})(\boldsymbol{u},V_0[\boldsymbol{\xi}_\beta]\boldsymbol{u})} &= 4\sigma^2\sum_{\alpha=1}^{N}\frac{\bar{\boldsymbol{\xi}}_\alpha\bar{\boldsymbol{\xi}}_\beta^\top}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})} = 4\sigma^2\bar{\boldsymbol{M}}. \qquad (29)
\end{aligned}$$

Thus, $E[\Delta_1^*\boldsymbol{M}\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u}]$ is

$$\begin{aligned}
E[\Delta_1^*\boldsymbol{M}\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u}] &= 8\sigma^2\sum_{\alpha=1}^{N}\frac{(\bar{\boldsymbol{M}}^{-}V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u},\bar{\boldsymbol{M}}\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha)\bar{\boldsymbol{\xi}}_\alpha}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2} \\
= 8\sigma^2\sum_{\alpha=1}^{N}\frac{(V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u},\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{M}}\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha)\bar{\boldsymbol{\xi}}_\alpha}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2} &= 8\sigma^2\sum_{\alpha=1}^{N}\frac{(V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u},\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha)\bar{\boldsymbol{\xi}}_\alpha}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}. \qquad (30)
\end{aligned}$$

The expectation of $\Delta_2^*\boldsymbol{M}\boldsymbol{u}$ is

$$E[\Delta_2^*\boldsymbol{M}\boldsymbol{u}] = -2\sum_{\alpha=1}^{N}\frac{E[(\Delta_1\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})\bar{\boldsymbol{\xi}}_\alpha(\Delta\boldsymbol{\xi}_\alpha,\boldsymbol{u})]}{(\boldsymbol{u},V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}$$

$$= 2\sum_{\alpha=1}^{N} \frac{E[(\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})(\Delta\boldsymbol{\xi}_\alpha, \boldsymbol{u})\bar{\boldsymbol{\xi}}_\alpha]}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}$$

$$= 2\sum_{\alpha=1}^{N} \frac{(\bar{\boldsymbol{M}}^{-}V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u}, E[\Delta_1\boldsymbol{M}\boldsymbol{u}\Delta\boldsymbol{\xi}_\alpha^\top\boldsymbol{u}])\bar{\boldsymbol{\xi}}_\alpha}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}. \tag{31}$$

We can evaluate $E[\Delta_1\boldsymbol{M}\boldsymbol{u}\Delta\boldsymbol{\xi}_\alpha^\top\boldsymbol{u}]$ as follows:

$$E[\Delta_1\boldsymbol{M}\boldsymbol{u}\Delta\boldsymbol{\xi}_\alpha^\top\boldsymbol{u}] = E[\sum_{\beta=1}^{N} \frac{(\Delta\boldsymbol{\xi}_\beta, \boldsymbol{u})\bar{\boldsymbol{\xi}}_\beta\Delta\boldsymbol{\xi}_\alpha^\top\boldsymbol{u}}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\beta]\boldsymbol{u})}] = \sum_{\beta=1}^{N} \frac{\bar{\boldsymbol{\xi}}_\beta(\boldsymbol{u}, E[\Delta\boldsymbol{\xi}_\beta\Delta\boldsymbol{\xi}_\alpha^\top]\boldsymbol{u})}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\beta]\boldsymbol{u})}$$

$$= 4\sigma^2 \frac{\bar{\boldsymbol{\xi}}_\alpha(\boldsymbol{u}, V_0[\Delta\boldsymbol{\xi}_\alpha]\boldsymbol{u})}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})} = 4\sigma^2\bar{\boldsymbol{\xi}}_\alpha. \tag{32}$$

Thus, $E[\Delta_2^*\boldsymbol{M}\boldsymbol{u}]$ is

$$E[\Delta_2^*\boldsymbol{M}\boldsymbol{u}] = 8\sigma^2\sum_{\alpha=1}^{N} \frac{(\bar{\boldsymbol{M}}^{-}V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u}, \bar{\boldsymbol{\xi}}_\alpha)\bar{\boldsymbol{\xi}}_\alpha}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2} = 8\sigma^2\sum_{\alpha=1}^{N} \frac{(V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u}, \bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha)\bar{\boldsymbol{\xi}}_\alpha}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}. \tag{33}$$

The expectation of $\Delta_2\boldsymbol{L}$ is

$$E[\Delta_2\boldsymbol{L}] = E[\sum_{\alpha=1}^{N} \frac{(\Delta\boldsymbol{\xi}_\alpha, \boldsymbol{u})^2 V_0[\boldsymbol{\xi}_\alpha]}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}] = \sum_{\alpha=1}^{N} \frac{(\boldsymbol{u}, E[\Delta\boldsymbol{\xi}_\alpha\Delta\boldsymbol{\xi}_\alpha^\top]\boldsymbol{u})V_0[\boldsymbol{\xi}_\alpha]}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}$$

$$= 4\sigma^2\sum_{\alpha=1}^{N} \frac{V_0[\boldsymbol{\xi}_\alpha]}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})} = 4\sigma^2\boldsymbol{N}. \tag{34}$$

The expectation of $\|\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u}\|^2$ is

$$E[\|\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u}\|^2] = E[(\bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u}, \bar{\boldsymbol{M}}^{-}\Delta_1\boldsymbol{M}\boldsymbol{u})]$$

$$= E[(\sum_{\alpha=1}^{N} \frac{\bar{\boldsymbol{\xi}}_\alpha(\Delta\boldsymbol{\xi}_\alpha, \boldsymbol{u})}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}, (\bar{\boldsymbol{M}}^{-})^2\sum_{\beta=1}^{N} \frac{\bar{\boldsymbol{\xi}}_\beta(\Delta\boldsymbol{\xi}_\beta, \boldsymbol{u})}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\beta]\boldsymbol{u})})]$$

$$= \sum_{\alpha,\beta=1}^{N} \frac{(\boldsymbol{u}, E[\Delta\boldsymbol{\xi}_\alpha\Delta\boldsymbol{\xi}_\beta^\top]\boldsymbol{u})(\bar{\boldsymbol{\xi}}_\alpha, (\bar{\boldsymbol{M}}^{-})^2\bar{\boldsymbol{\xi}}_\beta)}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\beta]\boldsymbol{u})}$$

$$= 4\sigma^2\sum_{\alpha=1}^{N} \frac{(\bar{\boldsymbol{\xi}}_\alpha, (\bar{\boldsymbol{M}}^{-})^2\bar{\boldsymbol{\xi}}_\alpha)}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})} = 4\sigma^2\mathrm{tr}(\sum_{\alpha=1}^{N} \frac{\bar{\boldsymbol{\xi}}_\alpha\bar{\boldsymbol{\xi}}_\alpha^\top}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}(\bar{\boldsymbol{M}}^{-})^2)$$

$$= 4\sigma^2\mathrm{tr}(\bar{\boldsymbol{M}}(\bar{\boldsymbol{M}}^{-})^2) = 4\sigma^2\mathrm{tr}(\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{M}}\bar{\boldsymbol{M}}^{-}) = 4\sigma^2\mathrm{tr}(\bar{\boldsymbol{M}}^{-}). \tag{35}$$

From eqs. (25)∼(35), the bias of the second order error $\Delta_2\boldsymbol{u}$ of eq. (23) is

$$E[\Delta_2\boldsymbol{u}] = 4\sigma^2\Big[\sum_{\alpha=1}^{N} \frac{(\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha + (\bar{\boldsymbol{\xi}}_\alpha, \bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_\alpha)\bar{\boldsymbol{M}}^{-}V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u}}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})^2}$$

$$-\mathrm{tr}(\bar{\boldsymbol{M}}^{-})\boldsymbol{u}\Big] + O(\sigma^4). \tag{36}$$

6 Hyperaccuracy Correction

The above analysis implies that we can obtain a hyperaccurate estimator by subtracting the bias $E[\Delta_2 \boldsymbol{u}]$, or its estimate, from the ML estimator $\hat{\boldsymbol{u}}$. Since the term $\mathrm{tr}(\bar{\boldsymbol{M}}^-)\boldsymbol{u}$ is for adjusting $\hat{\boldsymbol{u}}$ to have unit norm (Fig. 1), we need not consider it if we normalize the solution in the end. So, we correct the ML estimator $\hat{\boldsymbol{u}}$ in the form

$$\tilde{\boldsymbol{u}} = N[\hat{\boldsymbol{u}} - \Delta_c \boldsymbol{u}], \tag{37}$$

where $N[\,\cdot\,]$ denotes normalization into unit norm. The correction term $\Delta_c \boldsymbol{u}$ is given by

$$\Delta_c \boldsymbol{u} = 4\hat{\sigma}^2 \sum_{\alpha=1}^{N} \frac{(\boldsymbol{M}^- \boldsymbol{\xi}_\alpha, V_0[\boldsymbol{\xi}_\alpha]\hat{\boldsymbol{u}})\boldsymbol{M}^- \boldsymbol{\xi}_\alpha + (\boldsymbol{\xi}_\alpha, \boldsymbol{M}^- \boldsymbol{\xi}_\alpha)\boldsymbol{M}^- V_0[\boldsymbol{\xi}_\alpha]\hat{\boldsymbol{u}}}{(\hat{\boldsymbol{u}}, V_0[\boldsymbol{\xi}_\alpha]\hat{\boldsymbol{u}})}, \tag{38}$$

which is obtained from eq. (36) by omitting $O(\sigma^4)$, replacing $\boldsymbol{u}$ by $\hat{\boldsymbol{u}}$, and replacing $\bar{\boldsymbol{M}}$ by $\boldsymbol{M}$ defined by $\{\boldsymbol{\xi}_\alpha\}$. The variance σ^2 in eq. (24) is estimated by

$$\hat{\sigma}^2 = \frac{(\hat{\boldsymbol{u}}, \boldsymbol{M}\hat{\boldsymbol{u}})}{4(N-5)}. \tag{39}$$

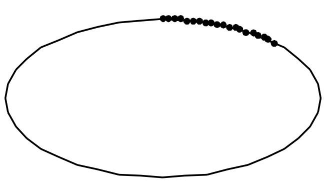
The approximations involved in eq. (38) may introduce errors of $O(\sigma)$ or higher, but they do not affect the leading order of eq. (38).

7 Experiments

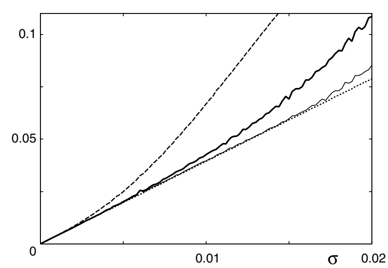
Fig. 2(a) shows $N = 20$ points $\{(\bar{x}_\alpha, \bar{y}_\alpha)\}$ taken on ellipse

$$\frac{x^2}{50^2} + \frac{y^2}{100^2} = 1 \tag{40}$$

with equal intervals. From them, we generated data points $\{(x_\alpha, y_\alpha)\}$ by adding Gaussian noise of mean 0 and standard deviation σ to the x and y coordinates



(a)



(b)

Fig. 2. (a) 20 points on an ellipse. (b) Noise level vs. fitting error: LS (broken line), ML (thick solid line), hyperaccuracy correction (thin solid line), KCR lower bound (dotted line).

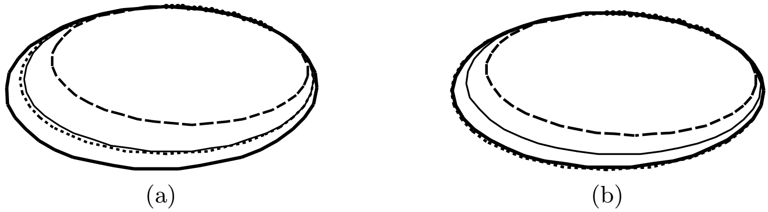


Fig. 3. Two instances of the fitted ellipse: LS (broken line), ML (thick solid line), hyperaccuracy correction (thin solid line), true ellipse (dotted line)

independently. Then, we fitted an ellipse by different methods. For computing ML, we used the FNS of Chojnacki et al. [7].

Fig. 2(b) plots for different σ the fitting error evaluated by the following root mean square over 10,000 independent trials:

$$E = \sqrt{\frac{1}{10000}\sum_{a=1}^{10000} \|\boldsymbol{P}_{\mathbf{u}}\hat{\boldsymbol{u}}^{(a)}\|^2}. \tag{41}$$

Here, $\hat{\boldsymbol{u}}^{(a)}$ is the ath value of $\hat{\boldsymbol{u}}$. Since its sign is indeterminate, we chose the one for which $(\hat{\boldsymbol{u}}^{(a)}, \boldsymbol{u}) \geq 0$. The thick solid line is for ML; the thin solid line is the result of our hyperaccurate correction. For comparison, we also plot the LS (least squares) solution $\hat{\boldsymbol{u}}_{\mathrm{LS}}$ by the broken line. The dotted line is the square root of the lower bound on $E[\|\boldsymbol{P}_{\mathbf{u}}\boldsymbol{u}\|^2]$ derived from eq. (7):

$$D = 2\sigma\sqrt{\mathrm{tr}\Big(\sum_{\alpha=1}^{N} \frac{\bar{\boldsymbol{\xi}}_\alpha \bar{\boldsymbol{\xi}}_\alpha^\top}{(\boldsymbol{u}, V_0[\boldsymbol{\xi}_\alpha]\boldsymbol{u})}\Big)^-}. \tag{42}$$

As can be seen from Fig. 2(b), the LS solution is not very accurate, while ML is very accurate; it almost coincides with the KCR lower bound when the noise is small. As the noise increases, however, a small gap appears between ML and the KCR lower bound. After adding the hyperaccurate correction, the accuracy almost coincides with the KCR lower bound.

Fig. 3(a) shows one instance of the fitted ellipse ($\sigma = 0.009$). The dotted line is the true ellipse; the broken line is for LS; the thick solid line is for ML; the thin solid line is for our hyperaccurate correction. We can see that the fitted ellipse is closer to the true shape after the correction. Fig. 3(b) is another instance (σ $= 0.009$). In this case, the ellipse given by ML is already very accurate, and it slightly deviates from the true shape after the correction.

Thus, the accuracy sometimes improves and sometimes deteriorates. Overall, however, the cases of improvement is the majority; on average we observe slight improvement as shown in Fig. 2(b). After close examination, we have observed that the accuracy drop occurs almost always when the ellipse fitted by ML falls inside the true shape. However, the majority of the fitted ellipses are outside the true shape. Thus, the correction is effective on average.

We infer that ML is likely to produce ellipses outside the true shape because it is parameterized in the form of eq. (1). If the major or minor axis of the ellipse is a, the coefficient of x^2 or y^2 is proportional to $1/a^2$. If $1/a^2$ is "unbiased", a is biased to be larger than the true value, as can be easily seen from the shape of the graph of $y = 1/x^2$.

8 Conclusions

We have demonstrated the existence of "hyperaccurate" ellipse fitting which outperforms ML. This is made possible by error analysis of ML followed by subtraction of high-order bias terms. However, ML nearly achieves the KCR lower bound, meaning that even if the bias is eliminated, the solution still fluctuates with the magnitude corresponding to the KCR lower bound, which is theoretically impossible to reduce. Thus, the accuracy improvement by our method is almost unnoticeable in practice, compared to which removing outliers and stabilizing the computation have far more practical significance. Nevertheless, our analysis has theoretical significance, illuminating the relationship between ML and the KCR lower bound.

Acknowledgments. The author thanks Nikolai Chernov of the University of Alabama, U.S.A., for helpful discussions. He also thanks Junpei Yamada of Okayama University, Japan, and Yasuyuki Sugaya of Toyohashi University of Technology, Japan, for doing numerical experiments. This work was supported in part by the Ministry of Education, Culture, Sports, Science and Technology, Japan, under a Grant in Aid for Scientific Research C (No. 17500112).

References

1. D. A. Anderson, The circular structural model, *J. R. Statist. Soc.*, B-**43**-2 (1981), 131–141.
2. M. Berman and D. Culpin, The statistical behaviour of some least squares estimators of the centre and radius of a circle, *J. R. Statist. Soc.*, B-**48**-2 (1986), 183–196.
3. F. J. Bookstein, Fitting conic sections to scattered data, *Comput. Graphics Image Process.*, **9** (1979) 56–71.
4. J. Cabrera and P. Meer, Unbiased estimation of ellipses by bootstrapping, *IEEE Trans. Patt. Mach. Intelli.*, **18**-7 (1996-7), 752–756.
5. N. N. Chan, On circular functional relationships, *J. R. Statist. Soc.*, B-**27** (1965), 45–56.
6. N. Chernov and C. Lesort, Statistical efficiency of curve fitting algorithms, *Comput. Stat. Data Anal.*, **47**-4 (2004-11), 713–728.
7. W. Chojnacki, M. J. Brooks, A. van den Hengel and D. Gawley, On the fitting of surfaces to data with covariances, *IEEE Trans. Patt. Anal. Mach. Intell.*, **22**-11 (2000-11), 1294–1303.
8. A. Fitzgibbon, M. Pilu and R. B. Fisher, Direct least square fitting of ellipses, *IEEE Trans. Patt. Anal. Mach. Intell.*, **21**-5 (1999-5), 476–480.

9. K. Kanatani, *Geometric Computation for Machine Vision*, Oxford University Press, Oxford, U.K., 1993.
10. K. Kanatani, Statistical bias of conic fitting and renormalization, *IEEE Trans. Patt. Anal. Mach. Intell.*, **16**-3 (1994-3), 320–326.
11. K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier Science, Amsterdam, The Netherlands, 1996, reprinted by Dover, New York, 2005.
12. K. Kanatani Cramer-Rao lower bounds for curve fitting, *Graphical Models Image Process.*, **60**-2, (1998-03), 93–99.
13. K. Kanatani, Further improving geometric fitting, *Proc. 5th Int. Conf. 3-D Digital Imaging and Modeling (3DIM 2005)*, June 2006, Ottawa, Ontario, Canada, pp. 2–13.
14. Y. Kanazawa and K. Kanatani, Optimal conic fitting and reliability evaluation, *IEICE Trans. Inf. & Sys.*, **E79-D**-9 (1996-9), 1323–1328.
15. D. Keren, D. Cooper and J. Subrahmonia, Describing complicated objects by implicit polynomials, *IEEE Trans. Patt. Anal. Mach. Intell.*, **16**-1 (1994-1), 38–53.
16. A. Kukush, I. Markovsky and S. Van Huffel, Consistent estimation in an implicit quadratic measurement error model, *Comput. Stat. Data Anal.*, **47**-1 (2004-8), 123–147.
17. A. Kukush and E. O. Maschlce, The efficiency of adjusted least squares in the linear functional relationship, *J. Multivariate Anal.*, **87-2** (2003-10), 261–274.
18. Y. Leedan and P. Meer, Heteroscedastic regression in computer vision: Problems with bilinear constraint, *Int. J. Comput. Vision.*, **37**-2 (2000), 127–150.
19. M. Mühlich and R. Mester, Unbiased errors-in-variables estimation using generalized eigensystem analysis, *Proc. 2nd Workshop Statistical Methods in Video Processing*, May 2004, Prague, Czech, pp. 38–49.
20. P. D. Sampson, Fitting conic sections to "very scattered" data: An iterative refinement of the Bookstein algorithm, *Comput. Graphics Image Process.*, **18** (1982), 97–108.
21. G. Taubin, Estimation of planar curves, surfaces, and non-planar space curves defined by implicit equations with applications to edge and rage image segmentation, *IEEE Trans. Patt. Anal. Mach. Intell.*, **13**-11 (1991-11), 1115–1138.
22. G. Taubin, F. Cukierman, S. Sullivan, J. Ponce and D. J. Kriegman, Parameterized families of polynomials for bounded algebraic curve and surface fitting, *IEEE Trans. Patt. Anal. Mach. Intell.*, **16**-3 (1994-3), 287–303.
23. M. Werman and D. Keren, A Bayesian method for filtering parametric and nonparametric models to noisy data, *IEEE Trans. Patt. Anal. Mach. Intell.*, **23**-5 (2001-5), 528–534.