Proc. Int. Conf. Patt. Recog. (ICPR 2010), August 2010, Istanbul, Turkey

Hyper Least Squares and Its Applications

Prasanna Rangarajan
Department of Electrical Engineering, Southern Methodist University, U.S.A.
prangara@mail.smu.edu

Kenichi Kanatani, Hirotaka Niitsuma
Department of Computer Science, Okayama University, Okayama 700-8530 Japan
{kanatani,niitsuma}@suri.cs.okayama-u.ac.jp

Yasuyuki Sugaya
Department of Computer Science and Engineering,
Toyohashi University of Technology, Toyohashi, Aichi 441-8580 Japan
sugaya@iim.cs.tut.ac.jp

Abstract

We present a new form of least squares (LS), called "hyperLS", for geometric problems that frequently appear in computer vision applications. Doing rigorous error analysis, we maximize the accuracy by introducing a normalization that eliminates statistical bias up to second order noise terms. Our method yields a solution comparable to maximum likelihood (ML) without iterations, even in large noise situations where ML computation fails.

1 Introduction

A fundamental problem in computer vision is the extraction of 2-D/3-D geometric information from noisy observations, for which the maximum likelihood (ML) estimator is known to provide a highly accurate solution [3, 4]. Unfortunately, ML computation is usually iterative and may not converge for high noise levels. It also requires an appropriate initial guess. The least squares (LS) estimator is a noniterative alternative to ML but is plagued by limited accuracy in the presence of noise. Doing rigorous error analysis, this paper presents a new LS estimator called "hyperLS" with accuracy comparable to ML. The improved accuracy results from introduction of a normalization that eliminates the statistical bias up to second order noise terms.

2 Geometric Fitting

Suppose noisy observations $\boldsymbol{x}_1$, ..., $\boldsymbol{x}_N$ are perturbations in the true values $\bar{\boldsymbol{x}}_1$, ..., $\bar{\boldsymbol{x}}_N$ that satisfy implicit geometric constraints of the form

$$F^{(k)}(\boldsymbol{x};\boldsymbol{\theta}) = 0, \qquad k = 1, ..., L. \tag{1}$$

The unknown parameter $\boldsymbol{\theta}$ allows us to infer the 2-D/3-D shape and motion of the observed objects [3]. Problems of this type are called *geometric fitting* [4]. In many important applications, we can reparameterize the problem to make the functions $F^{(k)}(\boldsymbol{x};\boldsymbol{\theta})$ linear in $\boldsymbol{\theta}$ (but nonlinear in $\boldsymbol{x}$) so that we can write Eq. (1) as

$$(\boldsymbol{\xi}^{(k)}(\boldsymbol{x}), \boldsymbol{\theta}) = 0, \qquad k = 1, ..., L, \tag{2}$$

where and hereafter $(\boldsymbol{a}, \boldsymbol{b})$ denotes the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. The vector $\boldsymbol{\xi}^{(k)}(\boldsymbol{x})$ represents a nonlinear mapping of $\boldsymbol{x}$.

Example 1. Given a point sequence (x_α, y_α), α = 1, ..., N, we wish to fit an ellipse of the form

$$Ax^2 + 2Bxy + Cy^2 + 2(Dx + Ey) + F = 0. \tag{3}$$

If we let

$$\boldsymbol{\xi} = (x^2\ 2xy\ y^2\ 2x\ 2y\ 1)^\top,\ \boldsymbol{\theta} = (A\ B\ C\ D\ E\ F)^\top, \tag{4}$$

Eq. (3) has the form of Eq. (2) for L = 1.

Example 2. Corresponding points (x, y) and (x', y') between two images of the same 3-D scene taken from different positions satisfy the *epipolar equation* [3]

$$(\boldsymbol{x}, \boldsymbol{F}\boldsymbol{x}') = 0,\ \ \boldsymbol{x} \equiv (x\, y\, 1)^\top,\ \ \boldsymbol{x}' \equiv (x'\, y'\, 1')^\top, \tag{5}$$

where $\boldsymbol{F}$ is called the *fundamental matrix*, from which we can compute the camera positions and the 3-D structure of the scene [3, 4]. If we let

$$\begin{aligned} \boldsymbol{\xi} &= (xx'\ xy'\ x\ yx'\ yy'\ y\ x'\ y'\ 1)^\top, \\ \boldsymbol{\theta} &= (F_{11}\ F_{12}\ F_{13}\ F_{21}\ F_{22}\ F_{23}\ F_{31}\ F_{32}\ F_{33})^\top, \end{aligned} \tag{6}$$

Eq. (5) has the form of Eq. (2) with L = 1.

Example 3. Two images of a planar or infinitely far away scene are related by a *homography* of the form

$$\boldsymbol{x}' \simeq \boldsymbol{H}\boldsymbol{x},\ \ \boldsymbol{x} \equiv (x\, y\, 1)^\top,\ \ \boldsymbol{x}' \equiv (x'\, y'\, 1')^\top, \tag{7}$$

where $\boldsymbol{H}$ is a nonsingular matrix, and $\simeq$ denotes equality up to nonzero multiplier [3, 4]. Equation (7) can alternatively be expressed as

$$\boldsymbol{x}' \times \boldsymbol{H}\boldsymbol{x} = \boldsymbol{0}. \tag{8}$$

If we let

$$\begin{aligned}
\boldsymbol{\xi}^{(1)} &= (0\ 0\ 0\ -x\ -y\ -1\ xy'\ yy'\ y')^\top, \\
\boldsymbol{\xi}^{(2)} &= (x\ y\ 1\ 0\ 0\ 0\ -xx'\ -yx'\ -x')^\top, \\
\boldsymbol{\xi}^{(3)} &= (-xy'\ -yy'\ -y'\ xx'\ yx'\ x'\ 0\ 0\ 0)^\top, \\
\boldsymbol{\theta} &= (H_{11}\ H_{12}\ H_{13}\ H_{21}\ H_{22}\ H_{23}\ H_{31}\ H_{32}\ H_{33})^\top,
\end{aligned} \tag{9}$$

The three component equations of Eq. (8) have the form of Eq. (2) for L = 3.

3 Algebraic Distance Minimization

For the sake of brevity, we abbreviate $\boldsymbol{\xi}^{(k)}(\boldsymbol{x}_\alpha)$ as $\boldsymbol{\xi}_\alpha^{(k)}$. *Algebraic* methods refer to those minimizing the *algebraic distance*

$$J = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k=1}^{L}(\boldsymbol{\xi}_\alpha^{(k)}, \boldsymbol{\theta})^2 = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k=1}^{L}\boldsymbol{\theta}^\top\boldsymbol{\xi}_\alpha^{(k)}\boldsymbol{\xi}_\alpha^{(k)\top}\boldsymbol{\theta} = (\boldsymbol{\theta}, \boldsymbol{M}\boldsymbol{\theta}), \tag{10}$$

where we define

$$\boldsymbol{M} = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k=1}^{L}\boldsymbol{\xi}_\alpha^{(k)}\boldsymbol{\xi}_\alpha^{(k)\top}. \tag{11}$$

Equation(10) is trivially minimized by $\boldsymbol{\theta} = \boldsymbol{0}$ unless scale normalization is imposed on $\boldsymbol{\theta}$. The most common normalization is $\|\boldsymbol{\theta}\| = 1$; we call this the *standard LS*. The crucial fact is that *the solution depends on the normalization*. The aim of this paper is to find a normalization that maximizes the accuracy of the solution. This issue has been raised by Al-Sharadqah and Chernov [1] and Rangarajan and Kanatani [9] for circle fitting, by Kanatani and Rangarajan [7] for ellipse fitting, and by Niitsuma et al. [8] for homography estimation. In this work, we generalize their results to an arbitrary number of constraints in Eq. (2). Following [1, 7, 8, 9], we consider the class of normalizations

$$(\boldsymbol{\theta}, \boldsymbol{N}\boldsymbol{\theta}) = \text{constant}. \tag{12}$$

Traditionally, the matrix $\boldsymbol{N}$ is assumed to be positive definite, but here we allow nondefinite (i.e., neither positive nor negative definite) matrices and search for $\boldsymbol{N}$ that maximizes the accuracy. If such an $\boldsymbol{N}$ is obtained, Eq. (10) is minimized subject to Eq. (12) by solving the generalized eigenvalue problem

$$\boldsymbol{M}\boldsymbol{\theta} = \lambda\boldsymbol{N}\boldsymbol{\theta}. \tag{13}$$

Evidently, λ = 0 in the absence of noise. If $\boldsymbol{N}$ is positive definite, the parameter $\boldsymbol{\theta}$ is estimated as the generalized eigenvector for the smallest eigenvalue λ, but in other cases for the smallest absolute value $|\lambda|$. Since the solution $\boldsymbol{\theta}$ of Eq. (13) has scale indeterminacy, we normalized it to $\|\boldsymbol{\theta}\| = 1$ rather than Eq. (12).

4 Error Analysis

Assuming that the noise $\Delta\boldsymbol{x}_\alpha$ in $\boldsymbol{x}_\alpha$ is independent and Gaussian with mean $\boldsymbol{0}$ and covariance matrix $V[\boldsymbol{x}_\alpha]$, we expand each $\boldsymbol{\xi}_\alpha^{(k)}$ in the form

$$\boldsymbol{\xi}_\alpha^{(k)} = \bar{\boldsymbol{\xi}}_\alpha^{(k)} + \Delta_1\boldsymbol{\xi}_\alpha^{(k)} + \Delta_2\boldsymbol{\xi}_\alpha^{(k)} + \cdots, \tag{14}$$

where $\bar{\boldsymbol{\xi}}_\alpha^{(k)}$ is the noiseless value, and $\Delta_i\boldsymbol{\xi}_\alpha^{(k)}$ is the ith order term in $\Delta\boldsymbol{x}_\alpha$. The first order term is written as

$$\Delta_1\boldsymbol{\xi}_\alpha^{(k)} = \boldsymbol{T}_\alpha^{(k)}\Delta\boldsymbol{x}_\alpha, \quad \boldsymbol{T}_\alpha^{(k)} \equiv \left.\frac{\partial\boldsymbol{\xi}^{(k)}(\boldsymbol{x})}{\partial\boldsymbol{x}}\right|_{\boldsymbol{x}=\bar{\boldsymbol{x}}_\alpha}. \tag{15}$$

We define the covariance of $\boldsymbol{\xi}_\alpha^{(k)}$, k = 1, ..., L, as follows ($E[\,\cdot\,]$ denotes expectation):

$$V^{(kl)}[\boldsymbol{\xi}_\alpha] \equiv E[\Delta_1\boldsymbol{\xi}_\alpha^{(k)}\Delta_1\boldsymbol{\xi}_\alpha^{(l)\top}] = \boldsymbol{T}_\alpha^{(k)}E[\Delta\boldsymbol{x}_\alpha\Delta\boldsymbol{x}_\alpha^\top]\boldsymbol{T}_\alpha^{(l)\top} = \boldsymbol{T}_\alpha^{(k)}V[\boldsymbol{x}_\alpha]\boldsymbol{T}_\alpha^{(l)\top}. \tag{16}$$

Substituting Eq. (14) into Eq. (11), we obtain

$$\boldsymbol{M} = \bar{\boldsymbol{M}} + \Delta_1\boldsymbol{M} + \Delta_2\boldsymbol{M} + \cdots, \tag{17}$$

where

$$\begin{aligned}
\bar{\boldsymbol{M}} &= \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k=1}^{L}\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(k)\top}, \\
\Delta_1\boldsymbol{M} &= \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k=1}^{L}(\bar{\boldsymbol{\xi}}_\alpha^{(k)}\Delta_1\boldsymbol{\xi}_\alpha^{(k)\top} + \Delta_1\boldsymbol{\xi}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(k)\top}), \\
\Delta_2\boldsymbol{M} &= \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k=1}^{L}(\bar{\boldsymbol{\xi}}_\alpha^{(k)}\Delta_2\boldsymbol{\xi}_\alpha^{(k)\top} + \Delta_1\boldsymbol{\xi}_\alpha^{(k)}\Delta_1\boldsymbol{\xi}_\alpha^{(k)\top} \\
&\qquad + \Delta_2\boldsymbol{\xi}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(k)\top}).
\end{aligned} \tag{18}$$

We also expand the solution $\boldsymbol{\theta}$ and λ of Eq. (13) in the form

$$\boldsymbol{\theta} = \bar{\boldsymbol{\theta}} + \Delta_1\boldsymbol{\theta} + \Delta_2\boldsymbol{\theta} + \cdots, \quad \lambda = \bar{\lambda} + \Delta_1\lambda + \Delta_2\lambda + \cdots. \tag{19}$$

Substituting Eqs. (17) and (19) into Eq. (13), we have

$$(\bar{\boldsymbol{M}} + \Delta_1\boldsymbol{M} + \Delta_2\boldsymbol{M} + \cdots)(\bar{\boldsymbol{\theta}} + \Delta_1\boldsymbol{\theta} + \Delta_2\boldsymbol{\theta} + \cdots) = (\bar{\lambda} + \Delta_1\lambda + \Delta_2\lambda + \cdots)\boldsymbol{N}(\bar{\boldsymbol{\theta}} + \Delta_1\boldsymbol{\theta} + \Delta_2\boldsymbol{\theta} + \cdots). \tag{20}$$

Equating terms of the same order, we obtain

$$\bar{\boldsymbol{M}}\bar{\boldsymbol{\theta}} = \bar{\lambda}\boldsymbol{N}\bar{\boldsymbol{\theta}}, \tag{21}$$

$$\bar{\boldsymbol{M}}\Delta_1\boldsymbol{\theta} + \Delta_1\boldsymbol{M}\bar{\boldsymbol{\theta}} = \bar{\lambda}\boldsymbol{N}\Delta_1\boldsymbol{\theta} + \Delta_1\lambda\boldsymbol{N}\bar{\boldsymbol{\theta}}, \tag{22}$$

$$\bar{\boldsymbol{M}}\Delta_2\boldsymbol{\theta} + \Delta_1\boldsymbol{M}\Delta_1\boldsymbol{\theta} + \Delta_2\boldsymbol{M}\bar{\boldsymbol{\theta}} = \bar{\lambda}\boldsymbol{N}\Delta_2\boldsymbol{\theta} + \Delta_1\lambda\boldsymbol{N}\Delta_1\boldsymbol{\theta} + \Delta_2\lambda\boldsymbol{N}\bar{\boldsymbol{\theta}} \tag{23}$$

We have $\bar{\boldsymbol{M}}\bar{\boldsymbol{\theta}} = \boldsymbol{0}$ for the true values, so $\bar{\lambda}$ = 0. The second of Eqs. (18) implies $(\bar{\boldsymbol{\theta}}, \Delta_1\bar{\boldsymbol{M}}\bar{\boldsymbol{\theta}})$ = 0. Computing the inner product of Eq. (22) and $\bar{\boldsymbol{\theta}}$ on both sides, we see that $\Delta_1\lambda$ = 0. Multiplying Eq. (22) by the pseudoinverse $\bar{\boldsymbol{M}}^-$ of $\bar{\boldsymbol{M}}$ from left, we obtain

$$\Delta_1\boldsymbol{\theta} = -\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\bar{\boldsymbol{\theta}}. \tag{24}$$

Note that since $\bar{\boldsymbol{M}}\bar{\boldsymbol{\theta}} = \boldsymbol{0}$, the matrix $\bar{\boldsymbol{M}}^-\bar{\boldsymbol{M}}$ ($\equiv \boldsymbol{P}_{\bar{\boldsymbol{\theta}}}$) is the projection operator in the direction orthogonal to $\bar{\boldsymbol{\theta}}$. Also, equating the first order terms in the expansion $\|\bar{\boldsymbol{\theta}}+\Delta_1\boldsymbol{\theta}+\Delta_2\boldsymbol{\theta}+\cdots\|^2 = 1$ shows $(\bar{\boldsymbol{\theta}}, \Delta_1\boldsymbol{\theta}) = 0$ [6], hence $\boldsymbol{P}_{\bar{\boldsymbol{\theta}}}\Delta_1\boldsymbol{\theta} = \Delta_1\boldsymbol{\theta}$. Substituting Eq. (24) into Eq. (23) and computing its inner product with $\bar{\boldsymbol{\theta}}$ on both sides, we obtain

$$\Delta_2\lambda = \frac{(\bar{\boldsymbol{\theta}}, \Delta_2\boldsymbol{M}\bar{\boldsymbol{\theta}}) - (\bar{\boldsymbol{\theta}}, \Delta_1\boldsymbol{M}\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\bar{\boldsymbol{\theta}})}{(\bar{\boldsymbol{\theta}}, \boldsymbol{N}\bar{\boldsymbol{\theta}})} = \frac{(\bar{\boldsymbol{\theta}}, \boldsymbol{T}\bar{\boldsymbol{\theta}})}{(\bar{\boldsymbol{\theta}}, \boldsymbol{N}\bar{\boldsymbol{\theta}})}, \tag{25}$$

where we put

$$\boldsymbol{T} = \Delta_2\boldsymbol{M} - \Delta_1\boldsymbol{M}\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}. \tag{26}$$

Next, we consider the second order error $\Delta_2\boldsymbol{\theta}$. Since $\boldsymbol{\theta}$ is normalized to have unit norm, we are interested in the error component orthogonal to $\bar{\boldsymbol{\theta}}$. So, we consider

$$\Delta_2^\perp\boldsymbol{\theta} \equiv \boldsymbol{P}_{\bar{\boldsymbol{\theta}}}\Delta_2\boldsymbol{\theta}\ (= \bar{\boldsymbol{M}}^-\bar{\boldsymbol{M}}\Delta_2\boldsymbol{\theta}). \tag{27}$$

Multiplying Eq. (23) by $\bar{\boldsymbol{M}}^-$ from left and substituting Eq. (24), we obtain

$$\Delta_2^\perp\boldsymbol{\theta} = \Delta_2\lambda\bar{\boldsymbol{M}}^-\boldsymbol{N}\bar{\boldsymbol{\theta}} + \bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\bar{\boldsymbol{\theta}} - \bar{\boldsymbol{M}}^-\Delta_2\boldsymbol{M}\bar{\boldsymbol{\theta}} = \frac{(\bar{\boldsymbol{\theta}}, \boldsymbol{T}\bar{\boldsymbol{\theta}})}{(\bar{\boldsymbol{\theta}}, \boldsymbol{N}\bar{\boldsymbol{\theta}})}\bar{\boldsymbol{M}}^-\boldsymbol{N}\bar{\boldsymbol{\theta}} - \bar{\boldsymbol{M}}^-\boldsymbol{T}\bar{\boldsymbol{\theta}}. \tag{28}$$

5 Covariance and Bias

The important observation is that the first order error $\Delta_1\boldsymbol{\theta}$ in Eq. (24) does *not* contain the matrix $\boldsymbol{N}$. This means that the leading term of the covariance matrix $V[\boldsymbol{\theta}] = E[\Delta_1\boldsymbol{\theta}\Delta_1\boldsymbol{\theta}^\top] + \cdots$ does not contain $\boldsymbol{N}$. Thus, *all algebraic methods have the same covariance matrix in the leading order*. This leads us to focus on the bias. Equation (24) implies the first order bias $E[\Delta_1\boldsymbol{\theta}]$ is $\boldsymbol{0}$, so the leading bias is $E[\Delta_2^\perp\boldsymbol{\theta}]$. To evaluate this, we first consider $E[\boldsymbol{T}]$. From the third of Eqs. (18), we see that $E[\Delta_2\boldsymbol{M}]$ is

$$E[\Delta_2\boldsymbol{M}] = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k=1}^{L}\Big(\bar{\boldsymbol{\xi}}_\alpha^{(k)}E[\Delta_2\boldsymbol{\xi}_\alpha^{(k)}]^\top + E[\Delta_1\boldsymbol{\xi}_\alpha^{(k)}\Delta_1\boldsymbol{\xi}_\alpha^{(k)\top}] + E[\Delta_2\boldsymbol{\xi}_\alpha^{(k)}]\bar{\boldsymbol{\xi}}_\alpha^{(k)\top}\Big) = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k=1}^{L}\Big(V^{(kk)}[\boldsymbol{\xi}_\alpha] + 2\mathcal{S}[\bar{\boldsymbol{\xi}}_\alpha^{(k)}\boldsymbol{e}_\alpha^{(k)\top}]\Big), \tag{29}$$

where we have used Eq. (16) and defined

$$\boldsymbol{e}_\alpha^{(k)} \equiv E[\Delta_2\boldsymbol{\xi}_\alpha^{(k)}]. \tag{30}$$

The operator $\mathcal{S}[\,\cdot\,]$ denotes symmetrization ($\mathcal{S}[\boldsymbol{A}] = (\boldsymbol{A}+\boldsymbol{A}^\top)/2$). The expectation $E[\Delta_1\boldsymbol{M}\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}]$ has the following form (the derivation is omitted; a similar computation appears in [6]):

$$E[\Delta_1\boldsymbol{M}\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}] = \frac{1}{N^2}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{L}\Big(\mathrm{tr}[\bar{\boldsymbol{M}}^-V^{(kl)}[\boldsymbol{\xi}_\alpha]]\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(l)\top} + (\bar{\boldsymbol{\xi}}_\alpha^{(k)}, \bar{\boldsymbol{M}}^-\bar{\boldsymbol{\xi}}_\alpha^{(l)})V^{(kl)}[\boldsymbol{\xi}_\alpha] + 2\mathcal{S}[V^{(kl)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{M}}^-\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(l)\top}]\Big). \tag{31}$$

From Eqs. (29) and (31), the expectation of $\boldsymbol{T}$ is

$$E[\boldsymbol{T}] = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k=1}^{L}\Big(V^{(kk)}[\boldsymbol{\xi}_\alpha] + 2\mathcal{S}[\bar{\boldsymbol{\xi}}_\alpha^{(k)}\boldsymbol{e}_\alpha^{(k)\top}]\Big) - \frac{1}{N^2}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{L}\Big(\mathrm{tr}[\bar{\boldsymbol{M}}^-V^{(kl)}[\boldsymbol{\xi}_\alpha]]\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(l)\top} + (\bar{\boldsymbol{\xi}}_\alpha^{(k)}, \bar{\boldsymbol{M}}^-\bar{\boldsymbol{\xi}}_\alpha^{(l)})V^{(kl)}[\boldsymbol{\xi}_\alpha] + 2\mathcal{S}[V^{(kl)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{M}}^-\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(l)\top}]\Big). \tag{32}$$

6 HyperLS

We propose to let $\boldsymbol{N} = E[\boldsymbol{T}]$. It follows from Eq. (28) that

$$E[\Delta_2^\perp\boldsymbol{\theta}] = \bar{\boldsymbol{M}}^-\Big(\frac{(\bar{\boldsymbol{\theta}}, E[\boldsymbol{T}]\bar{\boldsymbol{\theta}})}{(\bar{\boldsymbol{\theta}}, \boldsymbol{N}\bar{\boldsymbol{\theta}})}\boldsymbol{N} - E[\boldsymbol{T}]\Big)\bar{\boldsymbol{\theta}} = \boldsymbol{0}. \tag{33}$$

Since the right-hand side of Eq. (32) contains the true values $\bar{\boldsymbol{\xi}}_\alpha$ and $\bar{\boldsymbol{M}}$, we replace $\bar{\boldsymbol{x}}_\alpha$ in their definition by the observation $\boldsymbol{x}_\alpha$. This does not affect the result, since the odd order noise terms have expectation 0 and hence the resulting error in $E[\Delta_2^\perp\boldsymbol{\theta}]$ is of 4th oder. Thus, the second order bias is *exactly* 0. We call this scheme *hyper least squares*[1], or *hyperLS* for short, after Al-Sharadqah and Chernov [1].

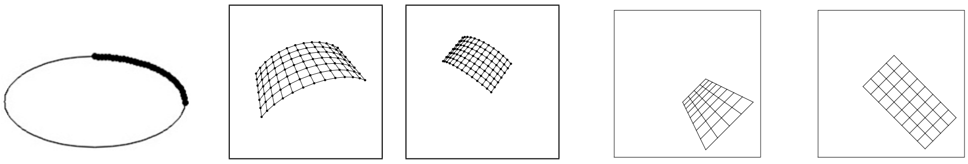
Note that $\boldsymbol{N}$ has scale indeterminacy: If $\boldsymbol{N}$ is multiplied by c ($\neq 0$), Eq. (13) has the same solution $\boldsymbol{\theta}$; only λ is divided by c. Thus, the noise characteristics $V^{(kl)}[\boldsymbol{\xi}_\alpha]$ in Eq. (16) and hence $V[\boldsymbol{x}_\alpha]$ need to be known only up to scale; *we need not know the absolute magnitude of the noise*.

Standard linear algebra routines for solving the generalized eigenvalue problem of Eq. (13) assume that $\boldsymbol{N}$ is positive definite, but here $\boldsymbol{N}$ is nondefinite. We circumvent this problem by rewriting (13) in the form

$$\boldsymbol{N}\boldsymbol{\theta} = \frac{1}{\lambda}\boldsymbol{M}\boldsymbol{\theta}. \tag{34}$$

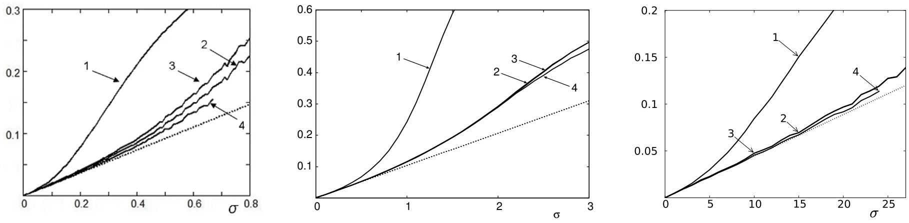
The matrix $\boldsymbol{M}$ in Eq. (11) is positive definite except in the absence of noise, in which case the smallest eigenvalue is 0. For large N, the second term $-(1/N^2)(\cdots)$ on the right-hand side of Eq. (32) has a smaller norm than the first term $(1/N)(\cdots)$, so it may be omitted. We call this the *Taubin approximation* from its similarity to the method due to Taubin [10].

[1]The origin of this term is Kanatani [5], who called his method "hyperaccurate", meaning it surpasses ML.



(a) (b) (c)

Figure 1. (a) 31 points on an ellipse. (b) Two views of a curved grid. (c) Two views of a planar grid.



(a) (b) (c)

Figure 2. RMS error vs. the standard deviation σ of noise added to each point. 1. standard LS, 2. hyperLS, 3. Taubin approximation, 4. ML. The dotted lines indicate the KCR lower bound. (a) Ellipse fitting. (b) Fundamental matrix computation. (c) Homography computation.

Example 4. We fit an ellipse to the point sequence shown in Fig. 1(a), compute the fundamental matrix between the two images shown in Fig. 1(b), and compute the homography relating the two images shown in Fig. 1(c). Independent Gaussian noise of mean 0 and standard deviation σ is added to the coordinates of each point. Figure 2 plots the RMS error of the computed parameter $\boldsymbol{\theta}$. The dotted lines indicate the theoretical accuracy limit called the *KCR lower bound* [4, 6]. In all examples, the standard LS performs poorly, while ML provides the highest accuracy (we used the method of Chojnacki et al. [2]). Note that ML computation fails in the presence of large noise (the plots interrupted in Fig. 2(a),(c)). In contrast, hyperLS can produce a solution close to ML in any noise level. For ellipse fitting, hyperLS is clearly superior to the Taubin approximation, while they are almost equivalent for fundamental matrices and homographies. This reflects that fact that while $\boldsymbol{\xi}$ is *quadratic* in x and y for ellipses (see Eqs. (4)), the corresponding $\boldsymbol{\xi}$ and $\boldsymbol{\xi}^{(k)}$ are *bilinear* in x, y, x', and y' for fundamental matrices (see Eqs. (6)) and homographies (see Eqs. (9)), so $\boldsymbol{e}_\alpha^{(k)}$ in Eq. (30) is $\boldsymbol{0}$.

7 Concluding Remarks

We presented a new form of least squares (LS), which we call "hyperLS", for geometric problems that frequently appear in computer vision applications. Doing rigorous error analysis, we maximized the accuracy by introducing a normalization that eliminates statistical bias up to second order noise terms. Numerical experiments of computing ellipses, fundamental matrices, and homographies show that our method yields a solution comparable to ML without iterations, even in large noise situations where ML computation fails.

References

[1] A. Al-Sharadqah and N. Chernov, Error analysis for circle fitting algorithms, *Elec. J. Stat.*, **3** (2009), 886–911.

[2] 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), 1294–1303.

[3] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, Cambridge, U.K., 2004.

[4] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier Science, Amsterdam, The Netherlands, 1996; reprinted, Dover, New York, 2005.

[5] K. Kanatani, Ellipse fitting with hyperaccuracy *IEICE Trans. Inf. & Syst.*, **E89-D**-10 (2006), 2653–2660.

[6] K. Kanatani, Statistical optimization for geometric fitting: Theoretical accuracy analysis and high order error analysis, *Int. J. Comput. Vis.*, **80**-2 (2008), 167–188.

[7] K. Kanatani and P. Rangarajan, Hyperaccurate algebraic ellipse fitting without iterations, *Proc. 5th Int. Conf. Comput. Vis. Theory Appl. (VISAPP'10)*, May 2010, Angers, France.

[8] H. Niistsuma, P. Rangarajan and K. Kanatani, High accuracy homography computation without iterations, *Proc. 16th Symp. Sensing Imaging Inf. (SSII'10)*, June 2010, Yokohama, Japan.

[9] P. Rangarajan and K. Kanatani, Improved algebraic methods for circle fitting, *Elec. J. Stat.*, **3** (2009), 1075–1082.

[10] G. Taubin, Estimation of planar curves, surfaces, and non-planar space curves defined by implicit equations with applications to edge and range image segmentation, *IEEE Trans. Patt. Anal. Mach. Intell.*, **13**-11 (1991), 1115–1138.