Improving the Error Estimation in Cosmological Parameter by the Third Order Expansion of the Natural Logarithm of the Likelihood

Jingyuan Chen
Advisor: Professor Dragan Huterer
Department of Physics, University of Michigan
450 Church Street, Ann Arbor, MI 48109

April 26, 2011

Abstract

The Fisher matrix formalism is a method of error forecasting and the method has been widely used in cosmological data analysis because of its convenience and effectiveness. It is based on the second-order expansion of the logarithm of the likelihood function in the parameter space, and the estimated error contours are ellipsoids in the parameter space. With real-world data, the error contours are usually have shapes more complicated than ellipsoids and hence the Fisher matrix formalism under- or overestimates errors. This paper investigates a generalization of the Fisher matrix formalism based on the third order expansion in the logarithm of the likelihood function.

1 Introduction

With the improvement in the precision in cosmological observations, cosmologists stared to work on improving the methodology in analyzing the data. Given some theoretical model, we want to estimate the model parameters via the observational data. This can be done using relevant equations given in the theoretical model. However, we would like to estimate the errors in the inferred parameters. This question has been investigated by the statistician R. A. Fisher, who developed the Fisher matrix formalism [1]. A brief and comprehensive review of the formalism can be found in [2]. The advantage of the Fisher matrix formalism is that it can be used to forecast the parameter errors and their covariances given the theoretical model without having the data itself. The formalism is convenient to be applied in numerical computations, and therefore it is currently widely used in cosmological parameter estimation. However, the Fisher matrix formalism has its limitations in under- or overestimating the errors in certain situations. In this paper we will try to investigate these misestimations.

2 Statistics Background

In Bayesian statistics, if we assume a null hypothesis Θ (including theoretical model and parameters) and have observations (e.g. data) $\mathbf{x}$, the Bayes Theorem says the conditional probabilities obey

$$p(\Theta|\mathbf{x})p(\mathbf{x}) = p(\mathbf{x}|\Theta)p(\Theta) \tag{1}$$

where $p(\mathbf{x})$ is the *evidence* based on the data and plays an important role in model selection [5]. $p(\Theta)$ is the *prior* probability based on the experimenter's knowledge (constant if no previous knowledge) [5]. $P(\mathbf{x}|\Theta)$ is called the *likelihood* and is usually denoted as $L(\mathbf{x};\Theta)$. $p(\Theta|\mathbf{x})$ is the *posterior* probability for the parameters given the observations $\mathbf{x}$.

In cosmology, we usually assume a fixed model and want to estimate the parameters θ (θ differs from Θ by that Θ specifies both the theoretical model and the parameters in it, while θ are just the parameters) and the errors in the estimation, therefore the evidence $p(\mathbf{x}) = \int_\theta p(\mathbf{x}|\theta)p(\theta)d\theta$, which is independent of θ, is simply an unimportant factor [3, 4, 5]. $\hat{\theta}$, defined as the best estimation of the parameters given the data, should occur when the posterior $p(\theta|\mathbf{x})$ is maximized. We often adopt the *principle of indifference*, assuming a constant prior, $p(\theta) \simeq const$ [3, 4, 5]. Thus, we have

$$p(\theta|\mathbf{x}) \propto L(\mathbf{x};\theta) \tag{2}$$

[3, 4, 5]. Therefore, at the the best estimation $\hat{\theta}$, we have $\partial_i L \equiv \partial L/\partial\theta^i = 0$ for each single parameter θ^i. We would also like to analyze the confidence level near $\hat{\theta}$.

The Fisher matrix formalism approximates the likelihood function in the parameter space by a Gaussian near the likelihood-maximizing parameters $\hat{\theta}$; equivalently it approximates $\ln L$ to the second order of $\delta\theta$ in Taylor series

$$\ln L(\hat{\theta} + \delta\theta) \simeq \ln L(\hat{\theta}) + \frac{\partial^2_{ij} \ln L(\hat{\theta})}{2!}\delta\theta^i\delta\theta^j \tag{3}$$

$$\Leftrightarrow \quad L(\hat{\theta} + \delta\theta) \simeq L(\hat{\theta}) \exp\left(\frac{\partial^2_{ij} \ln L(\hat{\theta})}{2!}\delta\theta^i\delta\theta^j\right) \tag{4}$$

(using the Einstein summation convention). Since the approximated likelihood function is Gaussian, given a certain confidence level, the error contour in the parameter space will be an n-dimensional ellipsoid (n being the number of parameters in θ), whose dimensions depend on the data $\mathbf{x}$. In the Fisher formalism, by assuming large data set, the dependence on particular data $\mathbf{x}$ converges to the dependence on only the mean μ and the covariance matrix $\mathbf{C}$, which can be calculated given the model and parameters without having the particular data. This makes the formalism useful in forecasting the expected error in an experiment or observation given its setup.

However, the actual likelihood function maybe far from a Gaussian. In such cases, if we do the Monte-Carlo simulation, the shape of the error contour will appear different from an ellipse, i.e., having "potato-like" or "banana-like" shapes. The Gaussian approximation will hence under- or overestimate the errors on certain directions in the parameter space [4]. In this paper we will explore some of the properties of the deformation from an ellipse to the actual shape of error contour by extending Fisher matrix formalism to the third order in the Taylor expansion in the parameter space.

3 The Third Order Approximation

Suppose we have a large data set $\mathbf{x}$, due to the Central Limit Theorem, the likelihood function will be Gaussian in the data space.

$$L \simeq \frac{1}{\sqrt{2\pi \det \mathbf{C}}} \exp\left(-\frac{(\mathbf{x}^T - \mu^T)\mathbf{C}^{-1}(\mathbf{x} - \mu)}{2}\right) \tag{5}$$

where μ is the prior mean and the $\mathbf{C}$ is the prior covariance matrix. Suppose we have a theoretical model Θ for the system, then μ and $\mathbf{C}$ are functions of the parameters θ via the equations given

in the model Θ. The Gaussian in the *data* space should not be confused with Fisher's Gaussian approximation in the *parameter* space. The Gaussian in the data space is a natural result of large data set and the Central Limit Theorem and requires no further assumption. The form of likelihood function in the parameter space, $L(\theta)$, however, depend on how $\mu, \mathbf{C}$ and θ are related, and hence is not a Gaussian in general; the Fisher matrix formalism approximates it by a Gaussian.

For most reasonable theoretical models, the likelihood function is well-behaved in the parameter space [3]. We can estimated the likelihood-maximizing parameters $\hat{\theta}$ by taking $\partial_i L(\hat{\theta}) = 0$ (notice that maximizing $\ln L$ is the same as maximizing L). It is more convenient to maximize $\ln L$ instead of L because L in related in Gaussian form (equation 5) to μ and $\mathbf{C}$ which are functions of θ. Expanding $\ln L$ in the parameter space near $\hat{\theta}$, where $\partial_i L(\hat{\theta}) = 0$, we have

$$\ln L(\hat{\theta} + \delta\theta) = \ln L(\hat{\theta}) + \frac{\partial^2_{ij} \ln L(\hat{\theta})}{2!}\delta\theta^i\delta\theta^j + \frac{\partial^3_{ijk} \ln L(\hat{\theta})}{3!}\delta\theta^i\delta\theta^j\delta\theta^k + ... \tag{6}$$

(using the Einstein summation convention).

As mentioned in the introduction (equation 4), the Fisher formalism takes the second order approximation in the above expression. The reason to do so is more than simply the ease of calculation [1]. Notice that $\partial^2_{ij} \ln L(\hat{\theta})$ in the second order coefficient is also Hessian matrix of $\ln L$ at $\hat{\theta}$. Since by definition $\hat{\theta}$ is a local maximum, the Hessian $\partial^2_{ij} \ln L(\hat{\theta})$ is negative definite (except for the special case of degeneracy where the Hessian is 0) and therefore equation 4 can be normalized. Any higher odd-order approximations diverge as $|\delta\theta| \to \pm\infty$ and hence can not be normalized; any higher even-order approximations may be normalizable but are not so in general. The likelihood function, which is a probability function, is supposed to be normalizable, therefore making the second order approximation a preferred choice.

In this paper we will investigate some properties of the third order approximation. The third order approximation is local but not normalizable. Since we are interested in the error in the estimation of θ, we are essentially considering the local region near $\hat{\theta}$ in the parameter space, and therefore a local approximation is sufficient.

The Fisher matrix is defined as

$$F_{ij} \equiv -\langle \partial^2_{ij} \ln L(\hat{\theta}) \rangle; \tag{7}$$

where the expectation $\langle\rangle$ is over data space. In this paper we will define a new quantity

$$G_{ijk} \equiv -\langle \partial^3_{ijk} \ln L(\hat{\theta}) \rangle. \tag{8}$$

We may approximate the likelihood function near $\hat{\theta}$ by

$$L(\hat{\theta} + \delta\theta) \simeq L(\hat{\theta}) \exp\left(-\frac{F_{ij}\delta\theta^i\delta\theta^j}{2} - \frac{G_{ijk}\delta\theta^i\delta\theta^j\delta\theta^k}{6}\right). \tag{9}$$

This is valid in the limit when the data set is large and hence $\overline{x} \to \mu, \overline{\mathbf{x}\mathbf{x}^T} - \mu\mu^T \to \mathbf{C}$. The dependence on data $\mathbf{x}$ can thus be absorbed to the dependence of only two quantities, μ, $\mathbf{C}$, and their derivatives which are prior given the theoretical model and parameters, as explained in Appendix A and [3, 4]. This allows us to forecast the error given the experiment setup and therefore we can expect the quality of the data from the experiment.

Notice that we cannot forecast $\hat{\theta}$ by similarly defining a quantity $\langle \partial_i L(\hat{\theta}) \rangle$ because it is also 0 as $\overline{\mathbf{x}} \to \mu$. In other words, we have to either have some prior knowledge about $\hat{\theta}$ or calculate $\hat{\theta}$ using the data $\mathbf{x}$.

From equation 5 we have $\ln L = -\frac{1}{2}\left[\ln(\det \mathbf{C}) + (\mathbf{x}^T - \mu^T)\mathbf{C}^{-1}(\mathbf{x} - \mu)\right] + const.$ By the definitions of μ and $\mathbf{C}$, we can derive (see Appendix A)

$$F_{ij} = \frac{1}{2}\mathbf{Tr}\left[\mathbf{C}^{-1}(\partial_i\mathbf{C})\mathbf{C}^{-1}(\partial_j\mathbf{C})\right] + \partial_i\mu^T\mathbf{C}^{-1}\partial_j\mu \tag{10}$$

and

$$\begin{aligned} G_{ijk} &= -\mathbf{Tr}\left[\mathbf{C}^{-1}(\partial_i\mathbf{C})\mathbf{C}^{-1}(\partial_j\mathbf{C})\mathbf{C}^{-1}(\partial_k\mathbf{C})\right] - \mathbf{Tr}\left[\mathbf{C}^{-1}(\partial_k\mathbf{C})\mathbf{C}^{-1}(\partial_j\mathbf{C})\mathbf{C}^{-1}(\partial_i\mathbf{C})\right] \\ &+\frac{1}{2}\mathbf{Tr}\left[\mathbf{C}^{-1}(\partial^2_{ij}\mathbf{C})\mathbf{C}^{-1}(\partial_k\mathbf{C})\right] + \frac{1}{2}\mathbf{Tr}\left[\mathbf{C}^{-1}(\partial^2_{ki}\mathbf{C})\mathbf{C}^{-1}(\partial_j\mathbf{C})\right] + \frac{1}{2}\mathbf{Tr}\left[\mathbf{C}^{-1}(\partial^2_{jk}\mathbf{C})\mathbf{C}^{-1}(\partial_i\mathbf{C})\right] \\ &-\partial_i\mu^T\mathbf{C}^{-1}(\partial_j\mathbf{C})\mathbf{C}^{-1}\partial_k\mu - \partial_k\mu^T\mathbf{C}^{-1}(\partial_i\mathbf{C})\mathbf{C}^{-1}\partial_j\mu - \partial_j\mu^T\mathbf{C}^{-1}(\partial_k\mathbf{C})\mathbf{C}^{-1}\partial_i\mu \\ &+\partial^2_{ij}\mu^T\mathbf{C}^{-1}\partial_k\mu + \partial^2_{jk}\mu^T\mathbf{C}^{-1}\partial_i\mu + \partial^2_{ki}\mu^T\mathbf{C}^{-1}\partial_j\mu. \end{aligned} \tag{11}$$

As we can see, F_{ij} and G_{ijk} are functions of μ and $\mathbf{C}$; $\mathbf{x}$ does not explicitly appear in them. Notice the symmetries

$$F_{ij} = F_{ji} \tag{12}$$

$$G_{ijk} = G_{jki} = G_{kij} = G_{ikj} = G_{jik} = G_{kji}. \tag{13}$$

Therefore for an n-dimensional parameter space we have $n(n-1)/2 + n$ independent entries in F_{ij} and $n(n-1)(n-2)/6 + n(n-1) + n$ in G_{ijk}.

4 Error Contour Deformation

In the Fisher formalism, an error contour is defined as [5]

$$F_{ij}\delta\theta^i\delta\theta^j = \chi^2 \tag{14}$$

for some constant χ^2. Along this contour,

$$L(\hat{\theta} + \delta\theta)/L(\hat{\theta}) = e^{const}. \tag{15}$$

Behind such definition of error contour in the Fisher formalism are the facts that for a Gaussian likelihood in the parameter space $\hat{\theta} = \overline{\theta}$, and the likelihood is symmetric about $\overline{\theta}$. However these facts do not hold for other likelihood functions in general. In such cases, since $\overline{\theta}$ is not of our interest, using the χ^2 contour for the definition of error contour near $\hat{\theta}$ is inconsistent as they have no association. Instead, we can define the error contour as *the contour along which equation 15 holds* for some *const.*; that is, we define the error contour by its relative likelihood compared to $\hat{\theta}$. In our third order case, the error contour should therefore have

$$F_{ij}\delta\theta^i\delta\theta^j + G_{ijk}\delta\theta^i\delta\theta^j\delta\theta^k/3 = const. \tag{16}$$

In cosmology and most other research, researchers usually present error contours in two dimensional parameter space spanned by some $\theta^{(1)}, \theta^{(2)}$ (if the likelihood function depend on more than two parameters, fix the other parameters at their $\hat{\theta^i}, i \neq 0, 1$). Define $\alpha \equiv \delta\theta^{(1)} = \theta^{(1)} - \hat{\theta}^{(1)}$ and $\beta \equiv \delta\theta^{(2)} = \theta^{(2)} - \hat{\theta}^{(2)}$ for convenience. For two parameters, F_{ij} has three independent entries $F_{\alpha\alpha}$, $F_{\alpha\beta}$ and $F_{\beta\beta}$, and G_{ijk} has four, $G_{\alpha\alpha\alpha}$, $G_{\alpha\alpha\beta}$, $G_{\alpha\beta\beta}$ and $G_{\beta\beta\beta}$.

The error contour in the Fisher formalism, given solely by F_{ij}, is an ellipse whose long and short axes can be found by the Lagrange multiplier method with the constraint $\alpha^2 + \beta^2 = const$

$$\nabla\left(F_{\alpha\alpha}\alpha^2 + 2F_{\alpha\beta}\alpha\beta + F_{\beta\beta}\right) \propto \nabla\left(\alpha^2 + \beta^2\right) \tag{17}$$

which yields

$$-F_{\alpha\beta}\alpha^2 + (F_{\alpha\alpha} - F_{\beta\beta})\alpha\beta + F_{\alpha\beta}\beta^2 = 0. \tag{18}$$

Defining the gradient $r \equiv \beta/\alpha$ (except for the trivial case $\alpha = 0$), we have

$$F_{\alpha\beta}(r^2 - 1) + (F_{\alpha\alpha} - F_{\beta\beta})r = 0 \tag{19}$$

which is a quadratic function from which we can solve for r, with the roots being the long and short axes of the ellipse. Notice that the quadratic discriminant is $(F_{\alpha\alpha} - F_{\beta\beta})^2 + 4F_{\alpha\beta}^2$ which is always nonnegative (equal to 0 only when $F_{\alpha\alpha} - F_{\beta\beta} = F_{\alpha\beta} = 0$, corresponding to a circle) and hence it always has roots.

Now we include the third order correction G_{ijk}, and the shape of the error contour deforms from an ellipse. To investigate the deformation, let's define

$$\begin{aligned} g(\alpha, \beta) &\equiv G_{ijk}\delta\theta^i\delta\theta^j\delta\theta^k \\ &= G_{\alpha\alpha\alpha}\alpha^3 + 3G_{\alpha\alpha\beta}\alpha^2\beta + 3G_{\alpha\beta\beta}\alpha\beta^2 + G_{\beta\beta\beta}\beta^3. \end{aligned} \tag{20}$$

Again if we take the gradient, g is a cubic function of r

$$g(r) = G_{\alpha\alpha\alpha} + 3G_{\alpha\alpha\beta}r + 3G_{\alpha\beta\beta}r^2 + G_{\beta\beta\beta}r^3 \tag{21}$$

and its cubic discriminant is

$$\begin{aligned} \Delta_g &= 18G_{\alpha\alpha\alpha}(3G_{\alpha\alpha\beta})(3G_{\alpha\beta\beta})G_{\beta\beta\beta} - 4(3G_{\alpha\alpha\beta})^3G_{\beta\beta\beta} + (3G_{\alpha\alpha\beta})^2(3G_{\alpha\beta\beta})^2 \\ &\quad -4G_{\alpha\alpha\alpha}(3G_{\alpha\beta\beta})^3 - 27G_{\alpha\alpha\alpha}^2G_{\beta\beta\beta}^2 \\ &= 162G_{\alpha\alpha\alpha}G_{\alpha\alpha\beta}G_{\alpha\beta\beta}G_{\beta\beta\beta} - 108(G_{\alpha\alpha\beta}^3G_{\beta\beta\beta} + G_{\alpha\alpha\alpha}G_{\alpha\beta\beta}^3) \\ &\quad +81G_{\alpha\alpha\beta}^2G_{\alpha\beta\beta}^2 - 27G_{\alpha\alpha\alpha}^2G_{\beta\beta\beta}^2 \end{aligned} \tag{22}$$

where

$$\begin{cases} \Delta_g > 0 & g \text{ has three distinct real roots} \\ \Delta_g = 0 & g \text{ has two distinct real roots, one of which is a multiple root} \\ \Delta_g < 0 & g \text{ has one real roots and two imaginary roots} \end{cases}.$$

We are going to discuss the behaviors of the deformation in each case.

Case I: $\Delta_g > 0$

There are three distinct real roots of $g(r)$, corresponding to three distinct lines which we call *zero axes*. On the other hand, *the principal axes* are the lines along which $g(r)$ are a local extremes with respect to r. Fixing the constraint $\alpha^2 + \beta^2 = const$ and use the Lagrange multiplier to find the extremes

$$\nabla g(\alpha, \beta) \propto \nabla(\alpha^2 + \beta^2) \tag{23}$$

the principal axes are given by the equation

$$G_{\alpha\alpha\alpha}\alpha^2\beta + 2G_{\alpha\alpha\beta}\alpha\beta^2 + G_{\alpha\beta\beta}\beta^3 = G_{\beta\beta\beta}\beta^2\alpha + 2G_{\alpha\beta\beta}\beta\alpha^2 + G_{\alpha\alpha\beta}\alpha^3 \tag{24}$$

$$G_{\alpha\beta\beta}r'^3 + (2G_{\alpha\alpha\beta} - G_{\beta\beta\beta})r'^2 - (2G_{\alpha\beta\beta} - G_{\alpha\alpha\alpha})r' - G_{\alpha\alpha\beta} = 0 \tag{25}$$

where $r' \equiv \beta/\alpha$ is the gradient of the principal axes (except for the trivial case $\alpha = 0$). The cubic discriminant of the above equation of principal axes, $\Delta_{p.a}$, can be shown that

$$\Delta_{p.a} > \Delta_g > 0 \tag{26}$$

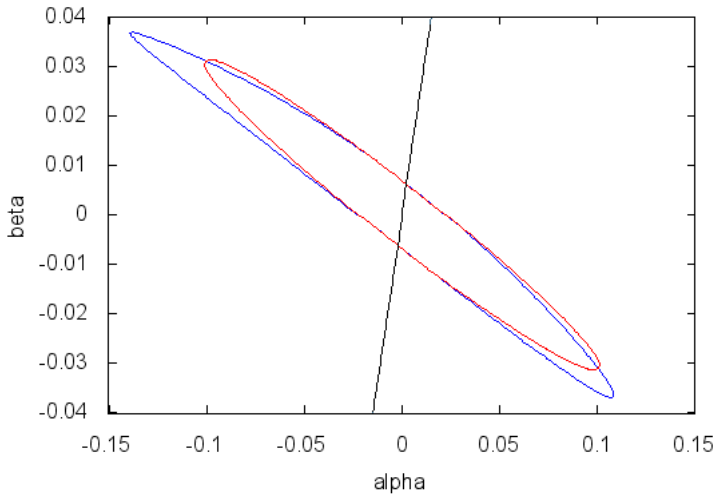

Figure 1: An example of the error contour bending (blue against red) with respect to a principal axis (black). The other two principal axes not shown. The F, G entries used are $F_{\alpha\alpha} = 1958, F_{\alpha\beta} = 6160, F_{\beta\beta} = 20394, G_{\alpha\alpha\alpha} = 4840, G_{\alpha\alpha\beta} = -6534, G_{\alpha\beta\beta} = -52602, G_{\beta\beta\beta} = -141900$. The gradient of the principal axis shown is 2.61.

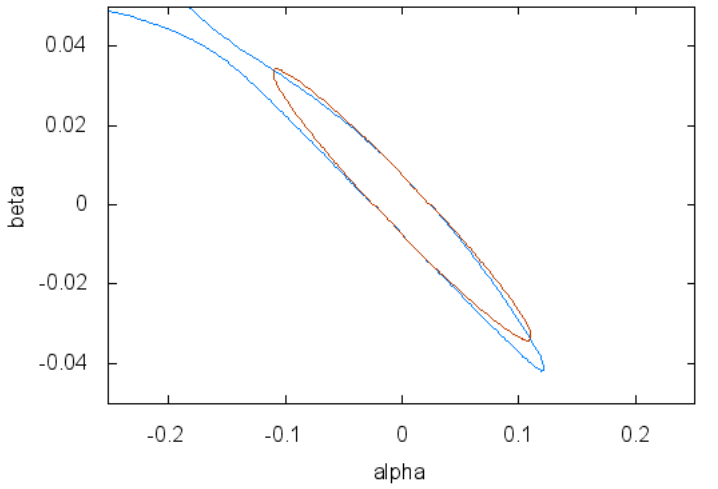

Figure 2: Same as Figure 1 except the entries here are 1.2 times larger. The divergence in the error contour is seen.

and therefore there are always three real roots of r', i.e. three principal axes, given $\Delta_g > 0$. Since g is a smooth function, the ordering of the three principal axes and the three zero axes must alternate as we increase the angle ϕ ($r = \tan\phi$) in the $\alpha\beta$-plane. For each principal axis, $g(\alpha, \beta)$ is positive on one end and negative on the opposite end, and between the positive ends of two principal axes there must be a negative end of the third principal axis.

Along a positive end, since $G_{ijk}\delta\theta^i\delta\theta^j\delta\theta^k > 0$, we need smaller error contour to satisfy equation 15, so the error contour shrinks on the positive end; along a negative end, the error contour extends out. The degree of deformation along an end depends on how close its two nearby zero axes are; if the two nearby zero axes are close, the deformation is large. If one of the principal axes happens to be almost parallel to the long axis of the Fisher ellipse, we can see the overall error contour bends toward the positive end of this principal axis and is thus deformed into a "banana-like" shape (Figure 1).

Notice that in Figure 2, in which all entries are 1.2 times larger than those in Figure 1, the error contour starts to diverge for certain large α, β; therefore we see the "tail" in the error contour. This is the due to the fact we mentioned in last section that any odd-approximation diverges for large $\delta\theta^i$. In fact any odd-order approximations have to diverge for large enough α and β along certain directions; the error contour in Figure 1, for example, diverges somewhere outside of the plot region, i.e. the diverging region is disconnected from the region in the the figure.

Case II: $\Delta_g = 0$

This case can simply be viewed as the limit of the case $\Delta_g > 0$ when two principle axes approach each other infinitely close.

Case III: $\Delta_g < 0$

Since g has only one root, there is only one zero axis. The deformation in this case is not dominated by the principal axes, for that $\Delta_{p.a}$ maybe positive, suggesting three distinct principal axes in contrast with only one zeros axis, not capturing the feature of g. Instead, it is dominated by the *protrusion axis* defined as the line along which

$$\nabla g(1, r') \cdot \langle 1, r\rangle = 0 \tag{27}$$

where r' is the gradient of the protrusion axis and r is the known gradient of the zero-axis. The meaning of the definition is, on any line in the phase space parallel to the zero axis, the extreme of g on this line is achieved at its intersection with the protrusion axis (see Figure 27). From equation 27 we can derive

$$r'^2(G_{\alpha\beta\beta} + rG_{\beta\beta\beta}) + 2r'(G_{\alpha\alpha\beta} + rG_{\alpha\beta\beta}) + (G_{\alpha\alpha\alpha} + rG_{\alpha\alpha\beta}) = 0 \tag{28}$$

which describes the relation between r' and r. Observe that $r' = r$ itself is a trivial solution to the equation; the other solution, i.e., the gradient of the protrusion axis, can therefore be solved using Viète's formula

$$r' = \frac{G_{\alpha\alpha\alpha} + rG_{\alpha\alpha\beta}}{r(G_{\alpha\beta\beta} + rG_{\beta\beta\beta})} = -\frac{2(G_{\alpha\alpha\beta} + rG_{\alpha\beta\beta})}{G_{\alpha\beta\beta} + rG_{\beta\beta\beta}} - r. \tag{29}$$

To better understand the essential property of the protrusion axis, we let $\langle a, c\rangle = \frac{1}{\sqrt{1+r^2}}\langle 1, r\rangle$ and $\langle b, d\rangle = \frac{1}{\sqrt{1+r'^2}}\langle 1, r'\rangle$. Consider the coordinate transformation

$$\begin{bmatrix} \alpha' \\ \beta' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{30}$$

after which the zero axis lies along $\langle \alpha', 0 \rangle$ and the protrusion axis along $\langle 0, \beta' \rangle$. We can define

$$g'(\alpha', \beta') \equiv G_{\alpha'\alpha'\alpha'}\alpha'^3 + 3G_{\alpha'\alpha'\beta'}\alpha'^2\beta' + 3G_{\alpha'\beta'\beta'}\alpha'\beta'^2 + G_{\beta'\beta'\beta'}\beta'^3 \tag{31}$$

and require it to preserve the invariance

$$g(\alpha, \beta) = g'(\alpha', \beta') \tag{32}$$

through the transformation from α, β to α', β'. Using this invariance, it can be derived that

$$G_{\alpha'\alpha'\alpha'} = 0 \tag{33}$$

$$G_{\alpha'\alpha'\beta'} = G_{\alpha\alpha\alpha}a^2b + G_{\beta\beta\beta}c^2d + G_{\alpha\alpha\beta}(a^2d + 2abc) + G_{\alpha\beta\beta}(bc^2 + 2acd) \tag{34}$$

$$G_{\alpha'\beta'\beta'} = 0 \tag{35}$$

$$G_{\beta'\beta'\beta'} = g(b, d) = G_{\alpha\alpha\alpha}b^3 + 3G_{\alpha\alpha\beta}b^2d + 3G_{\alpha\beta\beta}bd^2 + G_{\beta\beta\beta}d^3 \tag{36}$$

and hence

$$g'(\alpha', \beta') = \beta' \left(G_{\beta'\beta'\beta'}\beta'^2 + 3G_{\alpha'\alpha'\beta'}\alpha'^2 \right). \tag{37}$$

In equation 37 above, $\beta' = 0$ gives the zero axis as expected. On the other hand, for any line parallel to the zero axis, i.e. line with $\beta' = const.$, the extreme is at $\alpha' = 0$ which is along the protrusion axis as expected as well.

Similar to the principal axes in the $\Delta_g > 0$ case, the protrusion axis here also have a positive end and a negative end. As could be seen from equation 37, after the coordinate transformation, $\beta' G_{\alpha'\alpha'\beta'} > 0$ corresponds to the positive end and $\beta' G_{\alpha'\alpha'\beta'} < 0$ corresponds to the negative. However, in contrast to principal axes, the bending in this case is toward the negative end of the protrusion axis. This is due to the fact that, for $\beta' = const.$, the extreme $\alpha' = 0$ is a local minimum in terms of $|g'|$, therefore the deformation is the smallest along the protrusion axis and increases away from the protrusion axis.

The sharpness of the bending is defined as $3|G_{\alpha'\alpha'\beta'}/G_{\beta'\beta'\beta'}|$ (see equation 37), which is illustrated in Figure 3. Figure 4 illustrates that the larger the sharpness is, the more bending there is in the shape of the error contour. When the sharpness is 0, g' is reduced to proportional to β'^3 whose deforming is independent of α.

5 Application to a Supernova Survey

We use the supernovae survey as an example for illustrating the application of the third order correction in cosmology. Type Ia Supernovae (SNe Ia) has been used in the study of cosmological expansion since the late 1990's [6, 7]. A Type Ia Supernova is a result of an violent explosion of a white dwarf star, see [8] for details of the mechanism. SNe Ia produces a fixed peak luminosity (intrinsic luminosity) [6, 8], therefore it is used as a standard candle: by observing its flux, people can infer the luminosity distance (see below), providing information about the expansion history of the universe.

The luminosity distance is defined by the flux-luminosity relation [10, 11]

$$F = \frac{L}{A} = \frac{L}{4\pi d_L^2} \tag{38}$$

where F is the flux we observe, L is the intrinsic luminosity of the object (not to be confused with the likelihood function) which is fixed for SNe Ia, and d_L is the luminosity distance. The luminosity distance is related to cosmological parameters via [9, 10, 11]

$$d_L = (1 + z) \int_0^z \frac{dz'}{H(z')} \tag{39}$$

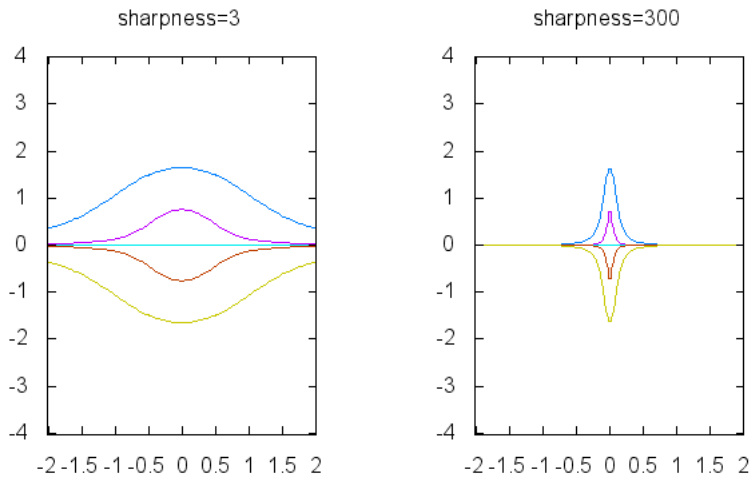

Figure 3: $g'(\alpha'\beta')$ (equations 37) contours for sharpness equal to 3 and 300, with $G_{\beta'\beta'\beta'} = 1/30$ fixed. The "protrusion" shape in the contours is along the protrusion axis. The horizontal axis is the zero axis and if we observe each of the horizontal lines, the extreme on it occur on the protrusion.

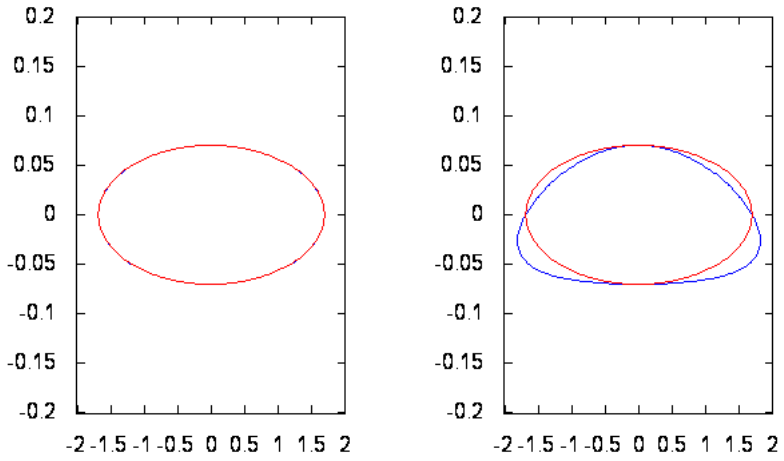

Figure 4: The error contours combining the protrusions in Figure 3 on an ellipse with equation $0.35\alpha'^2 + 100\beta'^2 = 0$. Large sharpness bends the ellipse more (the deformation on the left is so small to be seen); the direction of bending is along the protrusion axis which we see in Figure 3.

Table 1: F and G entries of 1000 supernovae survey between $z = 0$ and $z = 2$

F_{MM}	82713.21	G_{MMM}	-542813.33
F_{Mw}	24637.82	G_{MMw}	-204039.98
F_{ww}	7705.07	G_{Mww}	-24070.64
		G_{www}	19389.20

(M standing for Ω_M)

where z is the red-shift and H is the Hubble parameter. In a flat universe (curvature $k = 0$) with matter, dark energy and negligible radiation component, expanding the the Hubble parameter via the Friedman equations, we have

$$d_L(z) = (1+z)H_0^{-1} \int_0^z \frac{dz'}{\sqrt{\Omega_M(1+z')^3 + (1-\Omega_M)(1+z')^{3(1+w)}}}. \tag{40}$$

where Ω_M is the ratio of the density of matter to the critical density for a flat universe and $(1-\Omega_M)$ is thus the portion of dark energy; w is the equation-of-state of dark energy defined as $w = p_{DE}/\rho_{DE}$ which is the pressure-density ratio of dark energy. Current data in cosmology favor a flat universe dominated by matter (baryonic matter and dark matter) and dark energy, and therefore equation 40 is an effective description of the relation between d_L and the two cosmological parameters Ω_M and w.

In the SNe Ia survey the parameters to be estimated are Ω_M and w, using $\hat{\Omega}_M = 0.27$ and $\hat{w}_{DE} = -1$. The observable quantity is the distance modulus of the supernova, whose expectation is given by [6]

$$\mu(z) \equiv 5 \log_{10} d_L(z) + 25 \tag{41}$$

and is thus related to the parameters Ω_M, w via equation 40. Consider $\mu(z)$ for a sample of 1000 supernovae uniformly distributed from $z = 0$ to $z = 2$. The covariance matrix is conventionally taken to be constant $\mathbf{C}_{nm} = 0.15^2\delta_{nm}$ for supernova survey, where δ_{nm} is the Kronecker delta [12] and n, m takes 1 to 1000 (we use i, j, k to denote the indices of the parameters and m, n of the observable quantities).

Notice that, since $\mathbf{C}$ is constant, only the last term in F_{ij} (equation 10) and the last line in G_{ijk} (equation 11) are left, and i, j, k take Ω_M and w. Since the integrating variable z' in equation 40 is independent of Ω_M and w, we can take derivative of the integrand in equation 40 directly. The derivative of the integrand was performed explicitly and the integration over z' is carried out using Simpson's rule. The results are listed in Table 1.

The error contours corresponding to different $-2\ln(L(\theta)/L(\hat{\theta}))$ values (refer to equation 15) are provided in Figure 5. The size of the error contour decreases as we have larger data set, since we are thus able to do more precise estimation. Here we have 1000 supernovae, resulting in the error contour along which $-2\ln(L(\theta)/L(\hat{\theta})) = 1$ being very small. Deformation is seen in the $-2\ln(L(\theta)/L(\hat{\theta})) = 10$ and 100 contours. Notice that near the ends of the contour we see evidence of divergence, which was discussed in section 3.

Comparing with the real function error contours provided in sources [6, 11, 12], our third order approximation error contours in Figure 5 highly agree with them in shape, especially the direction of bending. The principal axis corresponding for direction of bending is $r' = 0.32$ in this case.

While Figure 5 and those real function error contours provided in sources mainly differ in three aspects.

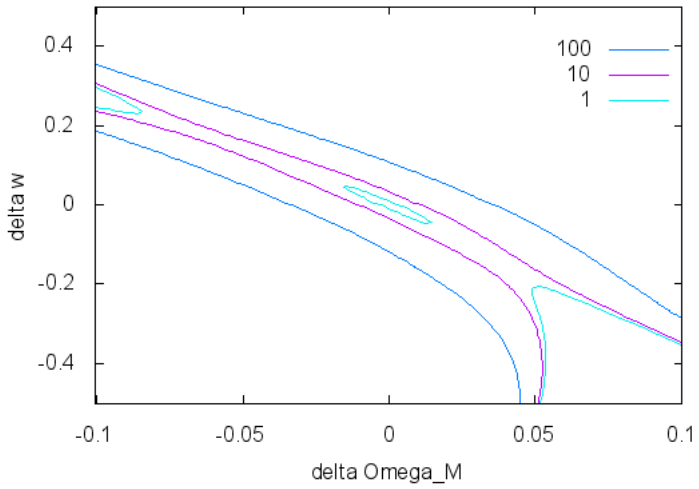

Figure 5: Error contours for a survey of 1000 supernovae between z=0 and 2 with $-2\ln(L(\theta)/L(\hat{\theta})) = 1, 10, 100$.

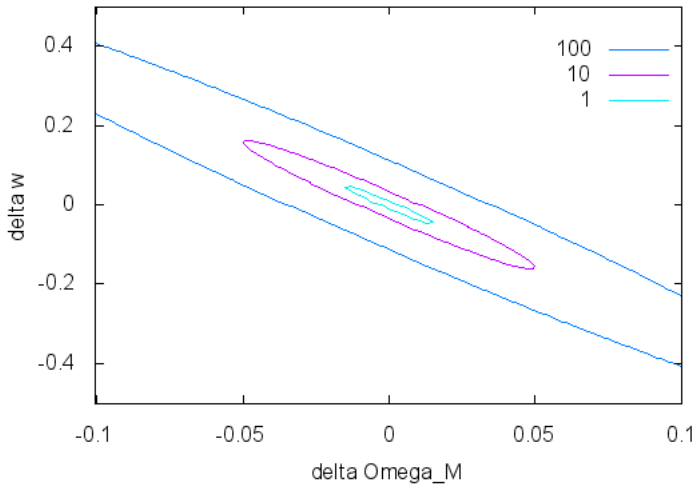

Figure 6: Error Contours estimated by the Fisher formalism (without G_{ijk})

1. The position of $\hat{\theta}$ are different (seen on the plots; the sources above did not explicitly provide the value of $\hat{\theta}$ they estimated). This is due to the fact discussed in the last part of section 3, that we have to provide knowledge about $\hat{\theta}$. The $\hat{\theta}$ we used was estimated using a combination of

SNe Ia survey, cosmic microwave background (CMB) and baryon acoustic oscillations (BAO) [11]; it is not surprising that it differs from the $\hat{\theta}$ estimated using solely SNe Ia survey. In summary this is a limitation originated from the Fisher matrix formalism, that we cannot assume $\langle \mathbf{x} \rangle \to \mu(\theta)$ and use it to forecast $\hat{\theta}$.

2. Our error contours diverge. This is an expected result as discussed in section 3. However this does not affect the effectiveness of our approximation capturing the local features of the real error contour. This is a particular limitation of our new third order correction and in general a limitation of any higher order corrections.

3. Comparing the plots, the size of our error contour is significantly smaller than the ones provided in the sources. This is not due to our third order correction; comparing Figure 6 with the real data plots, the same happen to the Fisher formalism without the third order correction. This is simply, as mentioned above, a result of the size set of data. This is not due to any particular limitation

These differences are all explainable and are expected.

6 Conclusion

We first reviewed the idea behind the Fisher matrix formalism. Its effectiveness comes from that by taking the large data set limit, the dependence of the likelihood on particular data $\mathbf{x}$ is absorbed into μ and $\mathbf{C}$. We then extend the same method to the third order of expansion in the parameter space in order to study the mis-estimation of error in the Fisher matrix formalism. In section 4 we discussed the mathematical properties of the new approximate term G_{ijk}, allowing us to forecast the deformation from an ellipse to the real error contour using simple quadratic and cubic algebras. Section 5 is an example illustrating the application of the new formalism in cosmological probes. The new formalism reproduced the deformation successfully.

The success of the Fisher matrix formalism is largely due to its convenience in computation. It is much faster than carrying out the Monte-Carlo simulation. The only intensive computational work in the Fisher formalism is the calculation of entries in F_ij; once it is calculated we can use it to forecast error with simple algebra. In our third order expansion formalism, the intensive computational work is in computing the G_{ijk} entries. Comparing equations 10 and 11, the calculations of entries in G_{ijk} are at least one order of magnitude slower than the calculation for F_{ij} (see Appendix B). This could be acceptable; yet any higher order approximations involve too much computational work to be used in research (see Appendix B).

The divergence of the third order approximation is undesirable since the likelihood function is probability function which is supposed to be normalized to one. One of the most significant consequence of this is the standard deviation is undefined, as the standard deviation is commonly used to define error. However we have to keep in mind that our estimation of θ is $\hat{\theta}$ which maximizes L, rather than the expectation value $\overline{\theta}$. As the standard deviation is associated with the expectation value by its definition, the use of standard deviation may not be suitable in our case; the use of equation 15 would therefore be a more reasonable definition.

In conclusion, in this paper we examined the extensions of the Fisher matrix error forecasting formalism and illustrated it on an example of synthetic SNe Ia data. In the future it would be good to further investigate the effectiveness of this new formalism with real data from various other cosmological probes such as cosmic microwave background (CMB), baryon acoustic oscillations (BAO) [11] and Galaxy clusters [13].

Appendix A: A Formalism of Determining Arbitrary Derivatives of $\ln L$

The derivation of F_{ij} and G_{ijk} starts with

$$\ln L = -\frac{1}{2}\left[\ln(\det \mathbf{C}) + (\mathbf{x}^T - \mu^T)\mathbf{C}^{-1}(\mathbf{x} - \mu)\right] + const \tag{42}$$

and uses the algebraic facts [4] that

$$\ln(\det \mathbf{C}) = Tr(\ln \mathbf{C}) \tag{43}$$

$$\partial_i(\mathbf{C}^{-1}) = -\mathbf{C}^{-1}(\partial_i \mathbf{C})\mathbf{C}^{-1}. \tag{44}$$

In general, consider the Nth derivative ($N > 1$), using the two formulae above and the derivative of multiplication, $d(uv) = u\ dv + v\ du$, we can prove by induction

$$\begin{aligned}
\partial^N_{i_1 \ldots i_N} \ln L \;=\; & \frac{1}{2} \sum_{[\phi]\in[\Phi_N]} (-1)^{|Im(\phi)|} \mathbf{Tr}\left\{\mathbf{C}^{-1}\left[\partial^{|\phi^{-1}(1)|}_{\phi^{-1}(1)}\mathbf{C}\right]\mathbf{C}^{-1}\left[\partial^{|\phi^{-1}(2)|}_{\phi^{-1}(2)}\mathbf{C}\right]\ldots\mathbf{C}^{-1}\left[\partial^{|\phi^{-1}(N)|}_{\phi^{-1}(N)}\mathbf{C}\right]\right\} \\
& -\frac{1}{2}\sum_{\psi\in\Psi_N}(-1)^{|Im(\psi)|}\left[\partial^{|\psi^{-1}(0)|}_{\psi^{-1}(0)}(\mathbf{x}^T - \mu^T)\right]\mathbf{C}^{-1}\left[\partial^{|\psi^{-1}(1)|}_{\psi^{-1}(1)}\mathbf{C}\right]\mathbf{C}^{-1}\left[\partial^{|\psi^{-1}(2)|}_{\psi^{-1}(2)}\mathbf{C}\right]\ldots \\
& \mathbf{C}^{-1}\left[\partial^{|\psi^{-1}(N)|}_{\psi^{-1}(N)}\mathbf{C}\right]\mathbf{C}^{-1}\left[\partial^{|\psi^{-1}(N+1)|}_{\psi^{-1}(N+1)}(\mathbf{x} - \mu)\right],
\end{aligned} \tag{45}$$

where the sets of functions Ψ_N and Φ_N are defined as

$$\Psi_N \equiv \{\psi \mid \psi : \{1, 2 \ldots N\} \to \{0, 1 \ldots N, N+1\}\} \tag{46}$$

$$\Phi_N \equiv \{\phi \mid \phi : \{1, 2 \ldots N\} \to \{1, 2 \ldots N\}\} \subset \Psi_N \tag{47}$$

and $[\Phi_N]$ is the quotient set of Φ_N up to cyclic permutations in the image of ϕ

$$[\Phi_N] \equiv \Phi_N / \text{ (cyclic permutations in the image of } \phi) \tag{48}$$

, namely each $[\phi] \in [\Phi_N]$ is an equivalence class in which every $\phi \in [\phi]$ is a cyclic permutations of the other elements in $[\phi]$. The set inverse is defined as

$$\psi^{-1}(n) = \{m \mid \psi(m) = n\} \tag{49}$$

, and $\partial^{|\psi^{-1}(n)|}_{\psi^{-1}(n)}$ means to take partial derivatives with respect to all i_j where $j \in \psi^{-1}(n)$ (if $\psi^{-1}(n) = \emptyset$ then no derivative is taken).

Now, we take the expectation of equation 45 with respect to $\mathbf{x}$. Notice that for $\psi \in \Phi_N$, $\psi^{-1}(0) = \psi^{-1}(N+1) = \emptyset$, hence, using $\mathbf{C} = \langle(\mathbf{x} - \mu)(\mathbf{x}^T - \mu^T)\rangle$ we have

$$\begin{aligned}
& \left\langle(\mathbf{x}^T - \mu^T)\mathbf{C}^{-1}\left[\partial^{|\psi^{-1}(1)|}_{\psi^{-1}(1)}\mathbf{C}\right]\ldots\mathbf{C}^{-1}\left[\partial^{|\psi^{-1}(N)|}_{\psi^{-1}(N)}\mathbf{C}\right]\mathbf{C}^{-1}(\mathbf{x} - \mu)\right\rangle \\
& = \mathbf{Tr}\left\{\mathbf{C}^{-1}\left[\partial^{|\psi^{-1}(1)|}_{\psi^{-1}(1)}\mathbf{C}\right]\ldots\mathbf{C}^{-1}\left[\partial^{|\psi^{-1}(N)|}_{\psi^{-1}(N)}\mathbf{C}\right]\right\}
\end{aligned} \tag{50}$$

(see [4]). Since traces are cyclic invariant, if $\psi_1 \in [\phi] \in [\Phi_N]$ and $\psi_2 \in [\phi] \in [\Phi_N]$ (notice, not in $[\Psi_N]$), then

$$\mathbf{Tr}\left\{\mathbf{C}^{-1}\left[\partial^{|\psi_1^{-1}(1)|}_{\psi_1^{-1}(1)}\mathbf{C}\right]\ldots\mathbf{C}^{-1}\left[\partial^{|\psi_1^{-1}(N)|}_{\psi_1^{-1}(N)}\mathbf{C}\right]\right\} = \mathbf{Tr}\left\{\mathbf{C}^{-1}\left[\partial^{|\psi_2^{-1}(1)|}_{\psi_2^{-1}(1)}\mathbf{C}\right]\ldots\mathbf{C}^{-1}\left[\partial^{|\psi_2^{-1}(N)|}_{\psi_2^{-1}(N)}\mathbf{C}\right]\right\} \tag{51}$$

, hence these terms can be combined to a coefficient equal to the size of $[\phi]$.

On the other hand, we have $\langle \mathbf{x} - \mu \rangle = 0$, therefore in taking the expectation of equation 45, the terms where $\psi^{-1}(0) = \emptyset \neq \psi^{-1}(N+1)$ or $\psi^{-1}(N+1) = \emptyset \neq \psi^{-1}(0)$ vanish.

Last, if $\psi_1^{-1}(0) = \psi_2^{-1}(N+1) \neq \emptyset$, $\psi_1^{-1}(N+1) = \psi_2^{-1}(0) \neq \emptyset$ and $\psi_1^{-1}(n) = \psi_2^{-1}(n)$ for other n's, then their corresponding terms are equal as C is symmetric. Thus, by enumerating all Ψ_2, Φ_2 and Ψ_3, Φ_3, we obtain the expressions for F and G in equations 10 and 11.

Appendix B: Computability

We can use the above formalism to estimate the computability of $\langle \partial^N_{i_1 \ldots i_N} \ln L \rangle$ with increasing N. First the cardinality of Ψ_N is $(N+2)^N$, which indicates that the number of terms in equation 45 is dominated by $(N+2)^N$. Second, the number of factors in each term in equation 45 is proportional to N. Last, consider the symmetries in taking the expectation. In the first line of equation 45, the number of symmetric terms increases proportionally to N; however, the symmetry in the second and third lines do not increase with N and thus these terms eventually dominate the computability. Therefore, the number of times of computation increase with N as $\sim N(N+2)^N$ which is so dramatic that any order higher than three will definitely not be of our interest.

However, notice that in the second and third lines in equation 45 most of the terms are intertwined terms of $\partial \mathbf{C}$ and $\partial \mu$. If one of $\mathbf{C}$ and μ is constant, then the complexity is simplified.

First consider the case of constant $\mathbf{C}$, like in the SNe Ia survey. Only terms of the form $\partial^n \mu \mathbf{C}^{-1} \partial^{N-n} \mu$ survive, where $1 \leq n \leq N-1$ (since $\langle \mathbf{x} - \mu \rangle = 0$, $n = 0$ or $n = N$ give zero terms). Thus is number of terms is given by $\frac{1}{2} \sum_{n=1}^{N-1} \frac{N!}{n!(N-n)!}$, where the factor $1/2$ is due to the fact that $\mathbf{C}$ is a symmetric matrix.

On the other hand, consider the case of constant μ, which is discussed in [3, 4] for Fisher matrix. In this case only terms in the first line in equation 45 survive. The cardinality of Φ is N^N and the cyclic symmetry is to the order of N. Therefore, with constant μ, the number of terms is two the order of N^{N-1}.

The typeset of the paper is by LaTeX; plots by gnuplot; numerical computations (C++) by Ms Visual Studio.

References

[1] R. A. Fisher, 1935, *The Logic of Inductive Inference*, J. Roy. Stat. Soc., 98, 39

[2] M. Tegmark, A. N. Taylor and A. F. Heavens, 1997, *Karhunen-Loève Eigenvalue Problems in Cosmology: How should we Tackle Large Data Sets?*, Astrophys. J., 480, 22

[3] S. Dodelson, 2003, *Modern Cosmology*, Academic Press, 1st ed.

[4] R. Durrer, 2008, *The Cosmic Microwave Background*, Cambridge University Press, 1st ed.

[5] A. F. Heavens, 2009, *Statistical Techniques in Cosmology*

[6] A. G. Reiss, et al., 1998, *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, Astron. J., 116, 1009

[7] G. Efstathiou, S. L. Bridle, A. N. Lasenby, M. P. Hobson and R. S. Ellis, 1999, *Constraints on Ω_λ and Ω_M from Distant Type Ia Supernovae and Cosmic Microwave Background Anisotropies*, Astron. Soc., 303, 47

[8] P. A. Mazzali, F. K. K. Röpke, S. Benetti and W. Hillebrandt, 2007, *A Common Explosion Mechanism for Type Ia Supernovae*, Science, 315, 825

[9] M. Li, X. D. Li and X. Zhang, 2010 *Comparison of dark energy models: A perspective from the latest observational data*, arXiv:0912.3988v3

[10] M. S. Turner and D. Huterer, 2007, *Cosmic Acceleration, Dark Energy, and Fundamental Physics*, J. Phys. Soc. Japan, 76, 11

[11] J. A. Frieman, M. S. Turner and D. Huterer, 2008, *Dark Energy and the Accelerating Universe*, Annu. Rev. Astro. Astrophys, 46, 385

[12] P. Astier, et al., 2006, *The Supernova Legacy Survey: Measurement of Ω_M, Ω_λ and w from the First Year Data Set*, A&A, 447, 31

[13] M. Tegmark, 1997, *Measuring Cosmological Parameters with Galaxy Surveys*, Phys. Rev., 79, 20