Lecture 10: Fundamentals of Observational Cosmology (IV)

Gong-Bo Zhao

Abstract

This lecture develops the Fisher matrix as a practical forecasting tool. Using supernova distances as the main example, we explain how to compute numerical derivatives, how parameter covariances become confidence ellipses, and why Fisher forecasts are fast and useful but not infallible.

Learning goals

After this lecture, students should be able to:

- define the Fisher matrix for a parameter-dependent model vector;
- estimate derivatives numerically using finite differences;
- interpret the covariance matrix of parameters;
- convert a 2×2 covariance matrix into an error ellipse;
- identify situations in which Fisher forecasts can become unreliable.

1 Fisher matrix for supernova distances

If the supernova covariance matrix is treated as fixed, then the Fisher matrix for parameters θ_i is

$$F_{ij} = \left(\frac{\partial \boldsymbol{\mu}}{\partial \theta_i}\right)^T \boldsymbol{\Sigma}^{-1} \left(\frac{\partial \boldsymbol{\mu}}{\partial \theta_j}\right),$$

where $\boldsymbol{\mu}$ is the vector of theoretical distance moduli evaluated at the observed redshifts.

Near the fiducial model, the parameter covariance is approximated by

$$\mathbf{C}_\theta \approx \mathbf{F}^{-1}.$$

This relation is the basis of most survey forecasts.

For SN Ia with parameters $(\Omega_{\rm m}, \Omega_\Lambda)$, the required ingredients are the numerical derivatives of the distance modulus with respect to each cosmological parameter,

$$\frac{\partial \mu}{\partial \Omega_{\rm m}}, \qquad \frac{\partial \mu}{\partial \Omega_\Lambda}.$$

In a realistic analysis, $\boldsymbol{\mu}$ is a vector with one entry for each supernova, so each derivative is itself a vector.

2 Numerical derivatives

In practice, derivatives are often evaluated by finite differences. The central-difference formula is

$$\frac{\partial f}{\partial x} \approx \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x}.$$

For a supernova analysis, this becomes

$$\frac{\partial \mu}{\partial \Omega_{\rm m}} \approx \frac{\mu(\Omega_{\rm m} + \Delta\Omega_{\rm m}, \Omega_\Lambda) - \mu(\Omega_{\rm m} - \Delta\Omega_{\rm m}, \Omega_\Lambda)}{2\Delta\Omega_{\rm m}},$$

and similarly for Ω_Λ,

$$\frac{\partial \mu}{\partial \Omega_\Lambda} \approx \frac{\mu(\Omega_{\rm m}, \Omega_\Lambda + \Delta\Omega_\Lambda) - \mu(\Omega_{\rm m}, \Omega_\Lambda - \Delta\Omega_\Lambda)}{2\Delta\Omega_\Lambda}.$$

Choice of step size. The derivative step must be chosen with care:

- if Δx is too large, the derivative is inaccurate because the function is not locally linear enough;
- if Δx is too small, round-off and numerical integration errors can dominate.

A practical habit is to test that the Fisher matrix is stable when the step size changes moderately.

3 From Fisher matrix to covariance matrix

For two parameters,

$$\mathbf{F} = \begin{pmatrix} F_{11} & F_{12} \\ F_{12} & F_{22} \end{pmatrix}, \qquad \mathbf{C}_\theta = \mathbf{F}^{-1}.$$

The diagonal entries of $\mathbf{C}_\theta$ are the variances:

$$\sigma_1^2 = (\mathbf{C}_\theta)_{11}, \qquad \sigma_2^2 = (\mathbf{C}_\theta)_{22}.$$

The off-diagonal term measures the covariance and therefore the tilt of the contour.

4 Error ellipses from a covariance matrix

Suppose the covariance matrix of two parameters x and y is

$$\mathbf{C} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}, \tag{1}$$

where ρ is the correlation coefficient. The corresponding quadratic form is

$$\Delta\chi^2 \equiv \mathbf{P}^T\mathbf{C}^{-1}\mathbf{P} = \frac{1}{1-\rho^2}\left(\frac{x^2}{\sigma_x^2} - 2\frac{\rho x y}{\sigma_x\sigma_y} + \frac{y^2}{\sigma_y^2}\right), \tag{2}$$

where $\mathbf{P} = (x, y)^T$ and the mean values have been subtracted.

At fixed $\Delta\chi^2$, this equation defines an ellipse. One way to draw the contour is to sample a fine grid in (x, y) and evaluate Eq. (2). That is mathematically transparent but not very efficient. For forecasting, it is usually faster to work directly with the parametric form of the ellipse.

4.1 Parametric form of the ellipse

A standard untilted ellipse can be written as

$$x'(t) = a\cos t, \qquad y'(t) = b\sin t, \qquad t \in [0, 2\pi],$$

where a and b are the semimajor and semiminor axes. If the ellipse is rotated counterclockwise by an angle θ, then

$$V = RV', \qquad V = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad V' = \begin{pmatrix} x' \\ y' \end{pmatrix}, \qquad R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{3}$$

Therefore,

$$x(t) = (a\cos\theta)\cos t - (b\sin\theta)\sin t, \tag{4}$$

$$y(t) = (a\sin\theta)\cos t + (b\cos\theta)\sin t. \tag{5}$$

So the problem reduces to determining a, b, and θ from the covariance matrix.

4.2 Rotation angle and principal axes

Under a rotation of coordinates,

$$\mathbf{C} = R\mathbf{C}'R^T, \qquad \mathbf{C}' = R^T\mathbf{C}R. \tag{6}$$

In the rotated frame, the ellipse is untilted, so $\mathbf{C}'$ is diagonal. This leads to

$$a^2 = \Delta\chi^2\, C'_{11} = \Delta\chi^2 \left(\sigma_x^2 \cos^2\theta + \rho\sigma_x\sigma_y \sin 2\theta + \sigma_y^2 \sin^2\theta\right), \tag{7}$$

$$b^2 = \Delta\chi^2\, C'_{22} = \Delta\chi^2 \left(\sigma_x^2 \sin^2\theta - \rho\sigma_x\sigma_y \sin 2\theta + \sigma_y^2 \cos^2\theta\right), \tag{8}$$

and the off-diagonal element must vanish. The condition $C'_{12} = 0$ gives

$$\tan 2\theta = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2}. \tag{9}$$

This is equivalent to the familiar formula in terms of the covariance C_{xy},

$$\tan 2\theta = \frac{2C_{xy}}{\sigma_x^2 - \sigma_y^2}.$$

An equivalent and often cleaner interpretation uses eigenvalues and eigenvectors:

- the eigenvectors of $\mathbf{C}$ define the principal directions of the ellipse;
- the eigenvalues λ_1 and λ_2 give the variances along those directions.

For a contour at fixed $\Delta\chi^2$,

$$a = \sqrt{\lambda_{\max}\,\Delta\chi^2}, \qquad b = \sqrt{\lambda_{\min}\,\Delta\chi^2}.$$

5 Confidence levels in two dimensions

For two jointly fitted parameters, the usual Gaussian confidence levels correspond to

$$\Delta\chi^2 \simeq 2.30, \qquad 6.18, \qquad 11.83$$

for 68.3%, 95.4%, and 99.7% confidence. Therefore a covariance matrix alone does not give the full contour until we choose the desired confidence level.

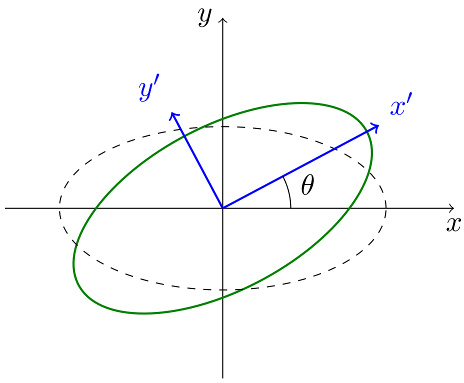

Figure 1: Illustration of an error ellipse. The solid green curve is a tilted contour in the original (x, y) coordinates. The dashed black ellipse represents the same contour in the rotated principal-axis coordinates (x', y'), where the covariance matrix is diagonal.

6 Combining independent datasets

If two datasets are independent, their log-likelihoods add, so their Fisher matrices also add:

$$\mathbf{F}_{\text{total}} = \mathbf{F}_1 + \mathbf{F}_2.$$

This is one of the main reasons Fisher forecasts are so popular in cosmology: combining projected constraints is fast and transparent.

7 Strengths and limitations of the Fisher approach

Strengths:

- very fast;
- easy to combine datasets;
- useful for survey design and parameter-degeneracy intuition.

Limitations:

- assumes the posterior is close to Gaussian near the fiducial point;
- can miss hard parameter boundaries and multimodality;
- depends on the chosen fiducial cosmology and derivative step sizes.

8 A practical workflow

A simple Fisher forecast proceeds as follows:

1. choose a fiducial cosmology;
2. compute the model vector $\boldsymbol{\mu}$;
3. evaluate numerical derivatives with respect to each parameter;
4. assemble the Fisher matrix;
5. invert it to obtain the parameter covariance and confidence ellipses.

9 Summary

The Fisher matrix is the local quadratic approximation to the likelihood around a fiducial model. It is one of the standard tools for cosmological forecasting because it is fast and interpretable. However, it should always be checked against full likelihood methods when non-Gaussian effects may be important.

Suggested reading

- Tegmark, Taylor, and Heavens (1997).
- Press, Teukolsky, Vetterling, and Flannery on numerical derivatives and finite-difference methods.
- Dodelson and Schmidt on Fisher methods and parameter forecasts.
- Forecast papers for supernovae, BAO, or weak lensing that visualize covariance ellipses.

Homework

1. **Central versus forward difference.** Write down both the forward-difference and central-difference approximations to $\partial f/\partial x$. Which is usually more accurate, and why?
2. **Combining datasets.** Starting from the fact that independent likelihoods multiply, show that independent Fisher matrices add.
3. **Ellipse geometry.** For the covariance matrix

$$\mathbf{C} = \begin{pmatrix} 4 & 1.5 \\ 1.5 & 1 \end{pmatrix},$$

compute the correlation coefficient and use the formula for $\tan 2\theta$ to determine the tilt angle.

4. **Eigenvalue interpretation.** Explain why the eigenvectors of the covariance matrix define the principal axes of the error ellipse.
5. **When Fisher fails.** Give two concrete situations in cosmology where a Fisher forecast could be misleading.
6. **Optional coding task.** Build a toy two-parameter Fisher matrix, invert it, and plot the corresponding 68% and 95% confidence ellipses.

References

[1] M. Tegmark, A. Taylor and A. Heavens, "Karhunen-Loève eigenvalue problems in cosmology: How should we tackle large data sets?," *Astrophysical Journal* **480**, 22 (1997), astro-ph/9603021.

[2] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in FORTRAN: The Art of Scientific Computing.* Cambridge University Press.