Solutions to the Homework in Lecture 10
Fundamentals of Observational Cosmology (IV)

These notes provide worked solutions to the homework problems in Lecture 10. We will repeatedly use the 2×2 covariance matrix

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

and the corresponding contour equation

$$\Delta\chi^2 = \boldsymbol{P}^T \mathbf{C}^{-1} \boldsymbol{P}.$$

Problem 1. Central versus forward difference

Question. Write down both the forward-difference and central-difference approximations to $\partial f/\partial x$. Which is usually more accurate, and why?

Solution. The **forward-difference** approximation is

$$\left.\frac{\partial f}{\partial x}\right|_x \approx \frac{f(x+h) - f(x)}{h}.$$

The **central-difference** approximation is

$$\left.\frac{\partial f}{\partial x}\right|_x \approx \frac{f(x+h) - f(x-h)}{2h}.$$

The central difference is usually more accurate. This follows from Taylor expansion. Around x,

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{6}f'''(x) + \mathcal{O}(h^4),$$

while

$$f(x-h) = f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{6}f'''(x) + \mathcal{O}(h^4).$$

For the forward difference,

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \frac{h}{2}f''(x) + \mathcal{O}(h^2),$$

so the leading truncation error is order h. For the central difference,

$$\frac{f(x+h) - f(x-h)}{2h} = f'(x) + \frac{h^2}{6}f'''(x) + \mathcal{O}(h^4),$$

so the leading truncation error is order h^2.

Therefore the central difference is normally preferred because its error decreases faster as the step size becomes small. The practical caveat, as emphasized in the lecture, is that if h is chosen *too* small, round-off or numerical-integration noise can still become a problem.

Problem 2. Combining datasets

Question. Starting from the fact that independent likelihoods multiply, show that independent Fisher matrices add.

Solution. Let two independent datasets have likelihoods $L_1(\boldsymbol{\theta})$ and $L_2(\boldsymbol{\theta})$. Independence implies

$$L_{\rm tot}(\boldsymbol{\theta}) = L_1(\boldsymbol{\theta})L_2(\boldsymbol{\theta}).$$

Taking the logarithm,

$$\ln L_{\rm tot} = \ln L_1 + \ln L_2.$$

Now differentiate twice with respect to the parameters θ_i and θ_j:

$$\frac{\partial^2 \ln L_{\rm tot}}{\partial\theta_i\partial\theta_j} = \frac{\partial^2 \ln L_1}{\partial\theta_i\partial\theta_j} + \frac{\partial^2 \ln L_2}{\partial\theta_i\partial\theta_j}.$$

By definition,

$$F_{ij} \equiv -\left\langle \frac{\partial^2 \ln L}{\partial\theta_i\partial\theta_j} \right\rangle.$$

Therefore

$$(F_{\rm tot})_{ij} = -\left\langle \frac{\partial^2 \ln L_{\rm tot}}{\partial\theta_i\partial\theta_j} \right\rangle = -\left\langle \frac{\partial^2 \ln L_1}{\partial\theta_i\partial\theta_j} \right\rangle - \left\langle \frac{\partial^2 \ln L_2}{\partial\theta_i\partial\theta_j} \right\rangle.$$

Hence

$$\boxed{\mathbf{F}_{\rm tot} = \mathbf{F}_1 + \mathbf{F}_2.}$$

This result is one of the main practical reasons Fisher methods are attractive: independent projected constraints can be combined by simple matrix addition.

Problem 3. Ellipse geometry

Question. 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.

Solution. From the matrix we identify

$$\sigma_x^2 = 4, \qquad \sigma_y^2 = 1, \qquad C_{xy} = 1.5.$$

So

$$\sigma_x = 2, \qquad \sigma_y = 1.$$

The correlation coefficient is

$$\rho = \frac{C_{xy}}{\sigma_x\sigma_y} = \frac{1.5}{2\times 1} = 0.75.$$

Therefore

$$\boxed{\rho = 0.75.}$$

Now use the rotation-angle formula

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

Substituting the numbers gives

$$\tan 2\theta = \frac{2(0.75)(2)(1)}{4-1} = \frac{3}{3} = 1.$$

Hence

$$2\theta = \arctan(1) = 45^\circ,$$

so one principal-axis direction is

$$\boxed{\theta = 22.5^\circ.}$$

The perpendicular principal axis is at

$$22.5^\circ + 90^\circ = 112.5^\circ.$$

Since the covariance is positive, the ellipse is tilted upward from left to right, consistent with a positive correlation between x and y.

Problem 4. Eigenvalue interpretation

Question. Explain why the eigenvectors of the covariance matrix define the principal axes of the error ellipse.

Solution. The error ellipse is defined by the quadratic form

$$\Delta\chi^2 = \boldsymbol{P}^T \mathbf{C}^{-1} \boldsymbol{P}, \qquad \boldsymbol{P} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Because the covariance matrix is real and symmetric, it can be diagonalized by an orthogonal matrix:

$$\mathbf{C} = R\Lambda R^T, \qquad \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

where the columns of R are the eigenvectors of $\mathbf{C}$ and λ_1, λ_2 are the eigenvalues. Introduce rotated coordinates

$$\mathbf{u} = R^T \boldsymbol{P}.$$

Then

$$\Delta\chi^2 = \mathbf{u}^T \Lambda^{-1} \mathbf{u} = \frac{u_1^2}{\lambda_1} + \frac{u_2^2}{\lambda_2}.$$

This is the equation of an *untilted* ellipse in the rotated coordinates u_1, u_2. So the directions that remove the tilt are precisely the eigenvector directions of the covariance matrix. That is why the eigenvectors define the principal axes.

The eigenvalues tell us how long the axes are. For a contour at fixed $\Delta\chi^2$,

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

So:

- **eigenvectors** $\rightarrow$ directions of the principal axes;
- **eigenvalues** $\rightarrow$ variances along those axes.

This is the cleanest mathematical interpretation of an error ellipse.

Problem 5. When Fisher fails

Question. Give two concrete situations in cosmology where a Fisher forecast could be misleading.

Solution. Two standard examples are the following.

1. Parameters near a physical boundary. Suppose we forecast a parameter that must be non-negative, such as the total neutrino mass $\sum m_\nu \geq 0$ or the tensor-to-scalar ratio $r \geq 0$. If the true value is close to zero, the real posterior is cut off by the boundary. A Fisher forecast, however, returns a symmetric Gaussian error bar centered on the fiducial value and therefore extends into the unphysical negative region. This makes the Gaussian approximation misleading.

2. Curved or strongly non-Gaussian degeneracies. Some cosmological parameter combinations produce banana-shaped or otherwise curved posteriors. Examples include combinations such as $(\Omega_m, \Omega_\Lambda)$ with limited geometric data or dark-energy parameters like (w_0, w_a) when the data constrain only certain redshift-weighted combinations. A Fisher matrix captures only the *local tangent ellipse* near the fiducial point. If the real contour bends strongly, the Fisher ellipse can underestimate the allowed region or misrepresent its orientation far from the best fit.

General lesson. Fisher methods are most reliable when the posterior is smooth, nearly Gaussian, and well inside parameter boundaries. They become less trustworthy when the posterior is truncated, multimodal, or strongly curved.

Problem 6. Optional coding task (Julia)

Question. Build a toy two-parameter Fisher matrix, invert it, and plot the corresponding 68% and 95% confidence ellipses.

Solution. Take a simple positive-definite Fisher matrix,

$$\mathbf{F} = \begin{pmatrix} 8 & -3 \\ -3 & 5 \end{pmatrix}.$$

Its inverse is the parameter covariance matrix,

$$\mathbf{C}_\theta = \mathbf{F}^{-1} = \frac{1}{31} \begin{pmatrix} 5 & 3 \\ 3 & 8 \end{pmatrix}.$$

Once we know the covariance matrix, we can diagonalize it, obtain the principal-axis directions and variances, and then draw the contours for

$$\Delta\chi^2 = 2.30 \quad \text{and} \quad 6.18,$$

which correspond to the usual 68% and 95% confidence regions in two dimensions.

The following Julia script performs all of these steps. The same code is also saved separately as `lecture10_fisher_ellipses.jl`.

```
using LinearAlgebra
using Plots

F = [8.0 -3.0;
     -3.0 5.0]

C = inv(F)
```

```
function ellipse_points(C, deltachi2; n = 400)
    eig = eigen(Symmetric(C))
    order = sortperm(eig.values; rev = true)
    vals = eig.values[order]
    vecs = eig.vectors[:, order]

    t = range(0, 2pi; length = n)
    circle = [cos.(t)'; sin.(t)']
    axes = Diagonal(sqrt.(vals .* deltachi2))
    pts = vecs * axes * circle
    return pts[1, :], pts[2, :]
end

x68, y68 = ellipse_points(C, 2.30)
x95, y95 = ellipse_points(C, 6.18)

plt = plot(x68, y68;
    aspect_ratio = :equal,
    xlabel = "parameter 1",
    ylabel = "parameter 2",
    linewidth = 2,
    label = "68% ellipse")
plot!(plt, x95, y95; linewidth = 2, linestyle = :dash, label = "95% ellipse")
scatter!(plt, [0.0], [0.0]; markersize = 4, label = "fiducial point")

savefig(plt, "lecture10_fisher_ellipses.pdf")
```

This script illustrates the standard Fisher workflow from the lecture: choose $\mathbf{F}$, invert it to obtain the covariance matrix, diagonalize the covariance, and use the eigenvalues and eigenvectors to draw the ellipses.