Project 1 - Part 1

Fisher Information Matrix

This notebook is arranged in cells. Texts are usually written in the markdown cells, and here you can use html tags (make it bold, italic, colored, etc). You can double click on this cell to see the formatting.

The ellipsis (...) are provided where you are expected to write your solution but feel free to change the template (not over much) in case this style is not to your taste.

Hit "Shift-Enter" on a code cell to evaluate it. Double click a Markdown cell to edit.

Write your partner's name here (if you have one).

Link Okpy

```
In [ ]:   1  from client.api.notebook import Notebook
          2  ok = Notebook('project1_p1_U.ok')
          3  _ = ok.auth(inline = True)
```

Imports

```
In [1]:   1  import numpy as np
          2  from scipy.integrate import quad
          3  #For plotting
          4  import matplotlib.pyplot as plt
          5  %matplotlib inline
```

Problem 1 - Constraining the cosmological parameters using the Planck power spectrum

Planck is the third-generation space telescope, following COBE and WMAP, and it aims to determine the geometry and content of the Universe by observing the cosmic microwave background radiation (CMB), emitted around 380,000 years after the Big Bang. Permeating the whole universe and containing information on the properties of the early Universe, the CMB is widely known as the strongest evidence for the Big Bang model.

Measuring the spectrum of the CMB, we confirm that it is very close to the radiation from an ideal blackbody, and flunctuations in the spectrum are very small. Averaging ocer all locations, its mean temperature is 2.725K, and its root mean square temperature fluctuation is $\langle(\frac{\delta T}{T})^2\rangle^{1/2} = 1.1 \times 10^{-5}$ (i.e. the temperature of the CMB varies by only ~ 30 μK across the sky).
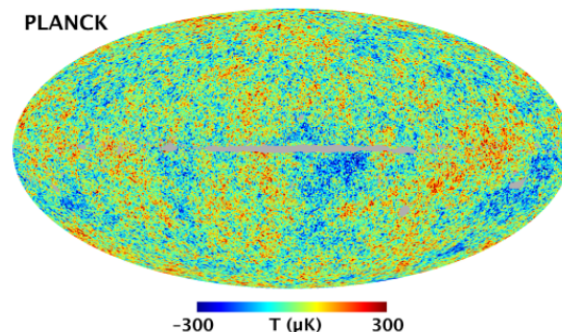
Suppose you observe the fluctuations $\delta T/T$. Since we are taking measurements on the surface of a sphere, it is useful to expand $\delta T/T$ in spherical harmonics (because they form a complete set of orthogonal functions on the sphere):

$$\frac{\delta T}{T}(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-1}^{l} a_{lm} Y_{lm}(\theta, \phi)$$

In flat space, we can do a Fourier transform of a function $f(x)$ as $\sum_k a_k e^{ikx}$ where k is the wavenumber, and $|a_k|$ determines the amplitude of the mode. For spherical harmonics, instead of k, we have l, the number of the modes along a meridian, and m, the number of modes along the equator. So l and m determine the wavelength ($\lambda = 2\pi/l$) and shape of the mode, respectively.

In cosmology, we are mostly interested in learning the statistical properties of this map and how different physical effects influence different physical scales, so it is useful to define the correlation function $C(\theta)$ and split the CMB map into different scales.

Suppose that we observe $\delta T/T$ at two different points on the sky. Relative to an observer, they are in direction $\hat{n}$ and $\hat{n}'$ and are separated by an angle θ given by $\cos\theta = \hat{n} \cdot \hat{n}'$ Then, we can find the correlation function by multiplying together the values of $\delta T/T$ at the two points and average the product over all points separated by the angle θ.

$$C(\theta)^{TT} = \left\langle \frac{\delta T}{T}(\hat{n})\frac{\delta T}{T}(\hat{n}') \right\rangle_{\hat{n} \cdot \hat{n}' = \cos\theta}$$

The above expression is specific to the temperature fluctuations, but we can also do a similar analysis for the polarization map of the CMB. (The CMB is polarized because it was scattered off of free electrons during decoupling.) We decompose the polarization pattern in the sky into a curl-free "E-mode" and grad-free "B-mode."

However, the CMB measurements (limited by the experiment resolution and the patch of sky examined) tell us about $C(\theta)$ over only a limited range of angular scales. (i.e. the precise values of $C(\theta)$ for all angles from $\theta = 0$ to $\theta = 180°$ is not known.) Hence, using the expansion of $\delta T/T$ in spherical harmonics, we write the correlation function as:

$$C(\theta) = \frac{1}{4\pi} \sum_{l=0}^{\infty} (2l + 1)C_l P_l(\cos\theta)$$

where P_l are the Legendre polynomials.

So we break down the correlation function into its multipole moments C_l, which is the angular power spectrum of the CMB.
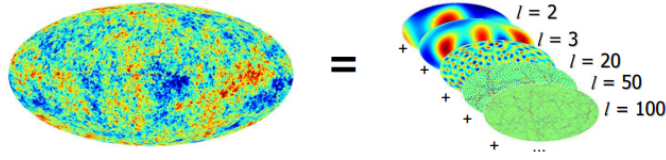
Remember that $\lambda = 2\pi / l$. So C_l measures the amplitude as a function of wavelength. ($C_l = \frac{1}{2l+1} \sum_{m=-l}^{l} |a_{lm}|^2$). In this problem, we will consider the temperature power spectrum C_l^{TT}, the E-mode power spectrum $C_l^{EE} = \frac{1}{2l+1} \sum_{m=-l}^{l} |a_{lm}^E|^2$, and the temperature-polarization cross-correlation $C_l^{TE} = \frac{1}{2l+1} \sum_{m=-l}^{l} a_{lm}^{T*} a_{lm}^E$.

THe CMB angular power spectrum is usually expressed in terms of $D_l = l(l+1)C_l/2\pi$ (in unit of μK^2) because this better shows the contribution toward the variance of the temperature fluctuations.

Cosmologists built a software called "cosmological boltzmann code" which computes the theoretical power spectrum given cosmological parameters, such as the Hubble constant and the baryon density. Therefore, we can fit the theory power spectrum to the measured one in order to obtain the best-fit parameters.

Here, we consider six selected cosmological parameters, $\vec{\theta} = [\theta_1, \theta_2, \ldots, \theta_6] = [H_0, \Omega_b h^2, \Omega_c h^2, n_s, A_s, \tau]$. ($H_0$ = Hubble constant, $\Omega_b h^2$ = physical baryon density parameter, $\Omega_c h^2$ = physical cold dark matter density parameter, n_s = scalar spectral index, A_s = curvature fluctuation amplitude, τ = reionization optical depth.). We provide you with the measured CMB power spectra from Planck Data Release 2.

References :
Intro to Cosmology, Barbara Ryden
http://folk.uio.no/hke/AST5220/v11/AST5220_2_2011.pdf (http://folk.uio.no/hke/AST5220/v11/AST5220_2_2011.pdf)
http://cosmology.berkeley.edu/~yuki/CMBpol/ (http://cosmology.berkeley.edu/~yuki/CMBpol/)

Problem - Fisher prediction for future CMB surveys

In class, we learned that the Fisher information matrix is useful for designing an experiment; we can vary the experiment design and predict the level of the expected error on any given parameter. In this problem, we aim to determine how well a low-noise, high-resolution future CMB survey would do in constraining the cosmological parameters.

The Fisher matrix is defined as the ensemble average of the Hessian of the log-likelihood ($\ln L$) with respect to the given parameters $\vec{\theta}$:

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

Here we take the model CMB power spectrum as our observables. (Here we consider the auto-correlations D_l^{TT}, D_l^{EE} and cross-correlation D_l^{TE} obtained from the boltzmann code using the best-fit cosmological parameters from Planck, https://arxiv.org/pdf/1502.01589v3.pdf (https://arxiv.org/pdf/1502.01589v3.pdf).) Then, we can estimate the Fisher matrix between two parameters θ_i and θ_j as:

$$F_{ij} = \sum_l \sum_k \frac{1}{(\sigma_l^k)^2} \frac{\partial D_l^k}{\partial \theta_i} \frac{\partial D_l^k}{\partial \theta_j}$$

where we sum over the CMB auto- and cross-power spectra $D_l^k = [D_l^0, D_l^1, D_l^2] = [D_l^{TT}, D_l^{EE}, D_l^{TE}]$, and we assume that there is no correlation between them. σ^2 is the variance of D_l and noise:

$$(\sigma_l^k)^2 = \frac{2}{(2l+1) \cdot f_{sky} \cdot \Delta l} (D_l^k + N_l^k)^2$$

In Part 1 and 2, we provide you with the model power spectra (D_l), its derivative with respect to the parameter θ ($\frac{\partial D_l}{\partial \theta}$), and the measurement error σ_l for the current Planck survey. In Part 3 and 4, you assume a zero-noise futuristic survey, so you need to compute new measurement error σ_l assuming $N_l = 0$.

f_{sky} is the fraction of the sky covered by the survey. Assume that $f_{sky} = 1$ for the sake of simplicity. Δl is the size of l-bin. Here, we set $l_{min} = 2$, $l_{max} = 2000$, and we have 92 l-bins in this range (For $2 \leq l < 30$, the power spectra are not binned ($\Delta l = 1$), and for $30 \leq l < 2000$, they are binned, and the bin size is $\Delta l = 30$). We obtain the measured and model power spectrum in that 92 l-bins. In this problem, first take the noise from Planck. We provide you with σ_l for D_l^{TT}, D_l^{EE}, D_l^{TE} from Planck.

Finally, we can obtain the covariance matrix C by inverting the Fisher matrix F:

$$[C] = [F]^{-1}$$

References:
Fisher Matrix Forecasting Review, Nicholas Kern
https://arxiv.org/pdf/0906.4123.pdf (https://arxiv.org/pdf/0906.4123.pdf)

1. First, load the measurement errors (σ_l^{TT}, σ_l^{EE}, σ_l^{TE}), model power spectrum (D_l^{TT}, D_l^{EE}, D_l^{TE}) and their derivatives with respect to six cosmological parameters evaluated at the best-fit values from Planck ($\frac{\partial D_l^{TT}}{\partial H_0}\big|_{\vec{\theta} = \vec{\theta}_{best-fit}}$, $\frac{\partial D_l^{TT}}{\partial \Omega_b h^2}\big|_{\vec{\theta} = \vec{\theta}_{best-fit}}$, etc). With the measurement errors from Planck, construct the Fisher matrix and the covariance matrix (you can use the numpy.linalg.inv for the matrix inversion). Evaluate the constraints on six parameters $\sigma(H_0)$, $\sigma(\Omega_b h^2)$, . . . , $\sigma(\tau)$ (corresponding to the square root of the diagonal entries of the covariance matrix). Print the results.

```
1  # Load data
2
3  # Best-fit values of the cosmological parameters from https://arxiv.org/pdf/1502.01589v3.pdf
4  H0      = 67.27
5  ombh2   = 0.02225
6  omch2   = 0.1198
7  ns      = 0.9645
8  As      = 2.2065e-9
9  tau     = 0.079
10
11 theta_best_Planck = np.array([H0, ombh2, omch2, ns, As, tau])
12
13
14 # Planck noise
15
16 # sigma_l for D_l^EE
17 data = np.loadtxt("Project1_EE_measured.dat")
18 # l (same for all power spectrum)
19 ell = data[:,0]
20 # and error
21 error_EE = data[:,2]
22
23 # sigma_l for D_l^TT
24 data = np.loadtxt("Project1_TT_measured.dat")
25 # and error
26 error_TT = data[:,2]
27
28 # sigma_l for D_l^TE
29 data = np.loadtxt("Project1_TE_measured.dat")
30 # and error
31 error_TE = data[:,2]
32
33
34 # Model power spectra given theta_best_Planck (calculated at the same ell bins as the measured power spectrum)
35
36 # D_l^EE (model)
37 data = np.loadtxt("Project1_EE_model_at_theta_best_Planck.dat")
38 EE_model_Planck = data[:,1]
39
40 # D_l^TT (model)
41 data = np.loadtxt("Project1_TT_model_at_theta_best_Planck.dat")
42 TT_model_Planck = data[:,1]
43
44 # D_l^TE (model)
45 data = np.loadtxt("Project1_TE_model_at_theta_best_Planck.dat")
46 TE_model_Planck = data[:,1]
47
48
49 # Derivative of the power spectrum given theta_best_Planck (calculated at the same ell bins as the measured power spectrum)
50
51 # Derivative of D_l^EE with respect to six parameters
52 # ([theta1, theta2, theta3, theta4, theta5, theta6] = [H_0, \Omega_b h^2, \Omega_c h^2, n_s, A_s, \tau])
53 data = np.loadtxt("Project1_derivative_EE_at_theta_best_Planck.dat")
54 deriv_DlEE_theta1 = data[:,1]
55 deriv_DlEE_theta2 = data[:,2]
56 deriv_DlEE_theta3 = data[:,3]
57 deriv_DlEE_theta4 = data[:,4]
58 deriv_DlEE_theta5 = data[:,5]
59 deriv_DlEE_theta6 = data[:,6]
60
61 # Derivative of D_l^TT with respect to six parameters
62 data = np.loadtxt("Project1_derivative_TT_at_theta_best_Planck.dat")
63 deriv_DlTT_theta1 = data[:,1]
64 deriv_DlTT_theta2 = data[:,2]
65 deriv_DlTT_theta3 = data[:,3]
66 deriv_DlTT_theta4 = data[:,4]
67 deriv_DlTT_theta5 = data[:,5]
68 deriv_DlTT_theta6 = data[:,6]
69
70 # Derivative of D_l^TE with respect to six parameters
71 data = np.loadtxt("Project1_derivative_TE_at_theta_best_Planck.dat")
72 deriv_DlTE_theta1 = data[:,1]
73 deriv_DlTE_theta2 = data[:,2]
74 deriv_DlTE_theta3 = data[:,3]
75 deriv_DlTE_theta4 = data[:,4]
76 deriv_DlTE_theta5 = data[:,5]
77 deriv_DlTE_theta6 = data[:,6]
```

In [ ]:
```
1  ...
```

In [ ]:
```
1  print('1-d constraints:')
2  print( 'H0 = %.2f +/- %.2f' % (..., ...), '\nOmega_b h^2 = %.5f +/- %.5f' % (..., ...), '\nOmega_c h^2 = %.4f +/- %.4f' % (..., ...), '\nn_s = %.
```

Now from the covariance matrix, we can plot 1-d and 2-d constraints on the parameters. (See Fig. 6 in Planck 2015 paper https://arxiv.org/pdf/1502.01589v3.pdf (https://arxiv.org/pdf/1502.01589v3.pdf))

1-d constraint (corresponding to the plots along the diagonal in Fig. 6, Planck 2015 paper):

First, the ith diagonal element of the covariance matrix correspond to $\sigma(\theta_i)^2$. Then, we can plot 1-d constraints on the parameter θ_i assuming a normal distribution with mean $= (\vec{\theta}_{\text{best}-\text{fit}})_i$ and variance $= \sigma(\theta_i)^2$.

2-d constraint (off-diagonal plots in Fig. 6, Planck 2015 paper):

Consider two parameters θ_i and θ_j from $\vec{\theta}$. Now marginalize over other parameters - in order to marginalize over other parameters, you can simply remove those parameters' row and column from the full covariance matrix. (i.e. From the full covariance matrix, you know the variance of all six parameters and their covariances with each other. So build a smaller dimension - 2 x 2 - covariance matrix from this.) - and obtain a 2×2 covariance matrix:

$$C_{ij} = \begin{pmatrix} \sigma(\theta_i)^2 & \text{Cov}(\theta_i, \theta_j) \\ \text{Cov}(\theta_i, \theta_j) & \sigma(\theta_j)^2 \end{pmatrix}$$

Now, we can plot the 2-dimensional confidence region ellipses from this matrix. The lengths of the ellipse axes are the square root of the eigenvalues of the covariance matrix, and we can calculate the counter-clockwise rotation of the ellipse with the rotation angle:

$$\tilde{\omega} = \frac{1}{2}\arctan\left(\frac{2 \cdot \text{Cov}(\theta_i, \theta_j)}{\sigma(\theta_i)^2 - \sigma(\theta_j)^2}\right) = \arctan\left(\frac{\vec{v_1}(y)}{\vec{v_1}(x)}\right)$$

where $\vec{v_1}$ is the eigenvector with the largest eigenvalue. So we calculate the angle of the largest eigenvector towards the x-axis to obtain the orientation of the ellipse.

Then, we multiply the axis lengths by some factor depending on the confidence level we are interested in. For 68%, this scale factor is $\sqrt{\Delta\chi^2} \approx 1.52$. For 95%, it is $\sqrt{\Delta\chi^2} \approx 2.48$.

Hint: For plotting ellipses, see HW4 Problem 1 Part 7.

2. From the covariance matrix, plot 1-d and 2-d constraints on the parameters. Note that the best-fit values of six parameters are alrady given in Part 1 (we just use the values from the Planck paper). For 2-d plot, show 68% and 95% confidence ellipses for pairs of parameters. You can arrange those subplots in a triangle shape, as in Fig. 6, Planck 2015 (https://arxiv.org/pdf/1502.01589v3.pdf (https://arxiv.org/pdf/1502.01589v3.pdf)).

```
In [ ]:   1  # Triangle Plot (Original skeleton code by Nicholas Kern)
          2  # Only a suggestion. You can create your own if you wish.
          3
          4  fig, axes = plt.subplots(6,6,figsize=(12,12))
          5  fig.subplots_adjust(wspace=0, hspace=0)
          6
          7  p_tex = np.array([r'$H_0$', r'$\Omega_bh^2$',r'$\Omega_ch^2$',r'$n_s$',r'$10^9 A_s$',r'$\tau$'])
          8
          9  for i in range(6):
         10      for j in range(6):
         11          ax = axes[i, j]
         12          if j > i:
         13              ax.axis('off')
         14              continue
         15          elif i == j:
         16              # diagonal part
         17              ax.grid(True)
         18              xarr = ...
         19              yarr = ...
         20              ax.plot(...)
         21              ax.set_xlim(...)
         22              ax.set_xticks(...)
         23              ax.set_yticklabels([])
         24              ax.set_xticklabels([])
         25          else:
         26              # off-diagonal part
         27              ax.grid(True)
         28
         29              # Covariance matrix
         30              CovM = ...
         31
         32              # Get eigenvalue/vector using svd
         33              eigvec, eigval, u = np.linalg.svd(CovM)
         34
         35              # Get Semimajor/minor axes of the ellipse
         36              semimaj = np.sqrt(eigval[0])*2.
         37              semimin = np.sqrt(eigval[1])*2.
         38
         39              # Rotation angle of the ellipse
         40              theta = np.arctan(eigvec[0][1]/eigvec[0][0])
         41
         42              # Create ellipses
         43              ell = mpl.patches.Ellipse(xy=[...], width=1.52*semimaj, height=1.52*semimin, angle = theta*180/np.pi, facecolor = 'dodgerblue', edged
         44              ell2 = mpl.patches.Ellipse(xy=[...], width=2.48*semimaj, height=2.48*semimin, angle = theta*180/np.pi, facecolor = 'skyblue', edgecol
         45
         46              ax.add_patch(ell2)
         47              ax.add_patch(ell)
         48
         49              ax.set_ylim(...)
         50              ax.set_xlim(...)
         51              ax.set_xticks(...)
         52              ax.set_yticks(...)
         53
         54
         55          if j != 0:
         56              ax.set_yticklabels([])
         57          if i != 5:
         58              ax.set_xticklabels([])
         59          if j == 0 and i !=0:
         60              ax.set_ylabel(p_tex[i], fontsize=10)
         61              ax.set_yticklabels(...)
         62              [tl.set_rotation(26) for tl in ax.get_yticklabels()]
         63          if i == 5:
         64              ax.set_xlabel(p_tex[j], fontsize=10)
         65              ax.set_xticklabels(...)
         66              [tl.set_rotation(26) for tl in ax.get_xticklabels()]
         67
```

Now, assume that we have an ideal, zero-noise CMB survey with $N_l = 0$. However, we are still instrinsically limited on the number of independent modes we can measure (there are only $(2l+1)$ of them)

- $C_l = \frac{1}{2l+1}\sum_{m=-l}^{l}\langle|a_{lm}|^2\rangle$. This leads that we get an instrinsic error (called "cosmic variance") in our estimate of C_l. So we approximate that

$$(\sigma_l^{EE})^2 = \frac{2}{(2l+1)\cdot f_{sky}\cdot\Delta l}(D_l^{EE})^2, \quad (\sigma_l^{TT})^2 = \frac{2}{(2l+1)\cdot f_{sky}\cdot\Delta l}(D_l^{TT})^2,$$

$$(\sigma_l^{TE})^2 = \frac{2}{(2l+1)\cdot f_{sky}\cdot\Delta l}\frac{(D_l^{TE})^2 + D_l^{TT}D_l^{EE}}{2}$$

.

3. Repeat Part 1 and 2 assuming $N_l^k = 0$. (How well does a zero-noise CMB survey constrain the cosmologial parameters?)

```
In [ ]:   1  ...
```

4. Combine Part 2 and Part 3 and compare. (First plot your results from Part 2 (1-d and 2-d constraints using the Planck power spectra and noise. Then, plot Part 3 results (assuming zero noise) on top with different colors. Note that your 1-d constrains in Part 3 are more sharply peaked Gaussians (with much smaller variances), so you can scale them so that its peak amplitudes match with your results from Part 2.)

```
In [ ]:    1 ...
```

5. Starting from the best-fit values from the Planck 2015 paper, you constrained six cosmological parameters assuming that you have a zero-noise future CMB survey. Compare your results with Table 3 and Figure 6 in https://arxiv.org/pdf/1502.01589v3.pdf (https://arxiv.org/pdf/1502.01589v3.pdf).

Answer:

To Submit

Execute the following cell to submit. If you make changes, execute the cell again to resubmit the final copy of the notebook, they do not get updated automatically.
We recommend that all the above cells should be executed (their output visible) in the notebook at the time of submission.
Only the final submission before the deadline will be graded.

```
In [ ]:    1 _ = ok.submit()
```