EEPS 1610 Solid Earth Geophysics:
In-Class Activity I- Seismic Tomography

Instructions

Work through **all four** of the sections **(section 4 is optional, and for extra credit)** and submit your final writeup, answering the questions in each section, via Canvas by Thursday, Oct. 21. Answer all the questions in *Italics*. Read and understand the **Background** section prior to beginning the activities. Be sure to refer to the file 'matlab_intro_f21.pdf' on canvas for suggestions and guidance on general MATLAB syntax.

Before you begin, you should download the following files from the EEPS 1610 Canvas site:

- DataFile_1610Lab1.mat
- Grid.mat

Learning Outcomes

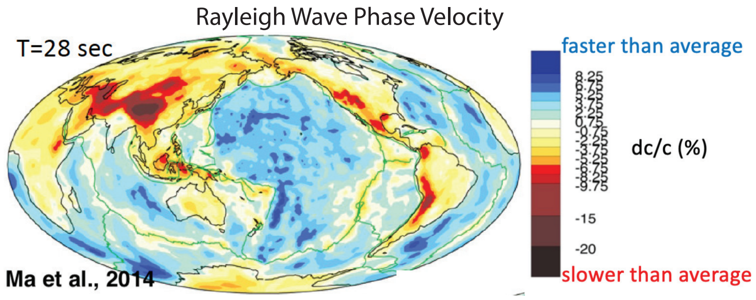
During this exercise, you will...

- Review the concepts of phase and group velocity, as well as phase and group traveltime measurements
- Gain experience plotting, processing, and generally working with large seismic datasets ($\approx 0.5 \times 10^6$ measurements) in MATLAB
- Understand the concept of seismic tomography and mathematics that underlie it
- Apply seismic tomography to image variations in Rayleigh wave phase velocity for a large swath of the U.S.
- Review the idea of the distributed depth sensitivity of Rayleigh wave measurements at a specific frequency, and gain experience interpreting seismic images in terms of the structure of the Earth's interior

Background

Motivation

Knowledge of the 3-D distribution of temperature, composition, and state throughout the Earth's interior would allow us to address fundamental questions involving our planet, including the nature and length-scales of mantle convection, the structure and dynamics of the lithosphere and asthenosphere, and the ultimate fate and life-cycle of subducting slabs. Since it is not possible to measure temperature, composition, and state directly, seismology offers the next best thing. Seismic tomography is a method that uses seismic waves to generate images of variations in seismic-wave speed inside the Earth. We can then convert these wave-speed images, with some assumptions, into images of temperature, composition, and state.



The image above is one that you saw during lecture. It is a global phase velocity map of the Earth's interior, and was constructed using hundreds of thousands of measurements. Today, you'll make an image just like this, but for a smaller region.

Mathematical Background

Our goal is to image a surface-wave velocity $c(x, y)$, which varies as a function of position—for example, longitude x and latitude y. To do this, we draw upon observations of the traveltimes of seismic waves traveling between two points (e.g., between an earthquake and a seismic station). The traveltime T of a wave traveling from point A to point B can be expressed in terms of the velocity as

$$T = \int_{s=A}^{s=B} \frac{1}{c(s)} ds$$

where the variable s represents position along the path between A and B. In other words, the total traveltime is related to the length of the path divided by the speed.

For example, the driving distance from Brown to Harvard is 53 miles (85 km). If you drive at 70 mph for 45 miles on the highway and at 30 mph for 8 miles on back roads, your total traveltime will be: 45miles/70mph + 8miles/30mph = 0.91 hours = 55 minutes.

Mathematically, we can write the integral above as a sum over N discrete segments of the path:

$$T = \sum_{i=1}^{N} L_i \frac{1}{c_i}$$

Where L_i is the distance traveled in the ith segment and c_i is the velocity traveled in that segment. Returning to the Brown-Harvard analogy, suppose your friends also drove from Brown to Harvard but took a different route, spending 30 miles on the highway (at 70 mph) and 23 miles on the back roads (at 30 mph). Their travel time will be longer—1.20 hours (72 minutes) instead of 0.91 hours. This yields a system of two equations:

$$T_A = \frac{L_{A1}}{c_1} + \frac{L_{A2}}{c_2} = \frac{45miles}{70mph} + \frac{8miles}{30mph} = 0.91hrs$$
$$T_B = \frac{L_{B1}}{c_1} + \frac{L_{B2}}{c_2} = \frac{30miles}{70mph} + \frac{23miles}{30mph} = 1.2hrs$$

where your time and lengths are indicated with the subscript A and your friends' are indicated with the subscript B. The highway speed and back-road speed are the same in both cases.

For imaging seismic-wave speeds in the Earth, the equation is the same but the objective is different: we do not know the speeds, but we do know the times and lengths. Let's pretend that's the case for the Brown-Harvard analogy:

$$\frac{45miles}{c_1} + \frac{8miles}{c_2} = 0.91hrs$$

$$\frac{30miles}{c_1} + \frac{23miles}{c_2} = 1.20hrs$$

There are two unknowns (c_1 and c_2), but there are also two equations, and therefore it is possible to solve for c_1 and c_2. In fact, it will be easiest if you define new variables that represent 1/c: $p_1 = 1/c_1$ and $p_2 = 1/c_2$. Then the equations become:

$$(45miles)(p_1) + (8miles)(p_2) = 0.91hrs$$

$$(30miles)(p_1) + (23miles)(p_2) = 1.20hrs$$

One way you may have been taught to solve a system of equations like this is to eliminate one variable. For example, rewrite the first equation as

$$p_1 = (0.91 - 8p_2)/45$$

Then the second equation becomes:

$$30(0.91 - 8p_2)/45 + 23p_2 = 1.20$$

and it is straightforward to solve for p_2 and then p_1.

However, when the system of equations has more than two unknowns, this quickly gets cumbersome. Matrix algebra (which is not required for this course) makes it so much easier. The system of equations can be written

$$\begin{pmatrix} 45miles & 8miles \\ 30miles & 23miles \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 0.91hours \\ 1.20hours \end{pmatrix}$$

In terms of mathematical notation, writing the two equations this way is identical to writing it as two separate equations. In matrix notation, we might write this as

$$\mathbf{Lp} = \mathbf{t}$$

where the matrix L contains the distances, the vector p contains the unknown p (=1/c) values, and the vector t contains the traveltimes. The beauty of using matrix algebra to solve for p is that a variety of methods exist to do so, even if there are many unknowns and therefore the size of L,p and t is large. MATLAB is incredibly good and fast at doing matrix algebra, even for large matrices.

In the context of our tomography problem in seismology, we can write that for k different traveltime measurements and a grid of N cells, we get:

$$\begin{pmatrix} L_{11} & L_{12} & ... & L_{1N} \\ L_{21} & L_{22} & ... & L_{2N} \\ .. & .. & ... & .. \\ L_{k1} & L_{k2} & ... & L_{kN} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ ... \\ p_N \end{pmatrix} = \begin{pmatrix} T_1 \\ T_2 \\ T_{...} \\ T_k \end{pmatrix}$$

Where L_{11} is the distance the earthquake corresponding to traveltime measurement 1 travels through cell 1, Where L_{12} is the distance the earthquake corresponding to traveltime measurement 1 travels through cell 2, and so on. We then solve for the slownesses in cells 1 through N by solving for the vector **p**.

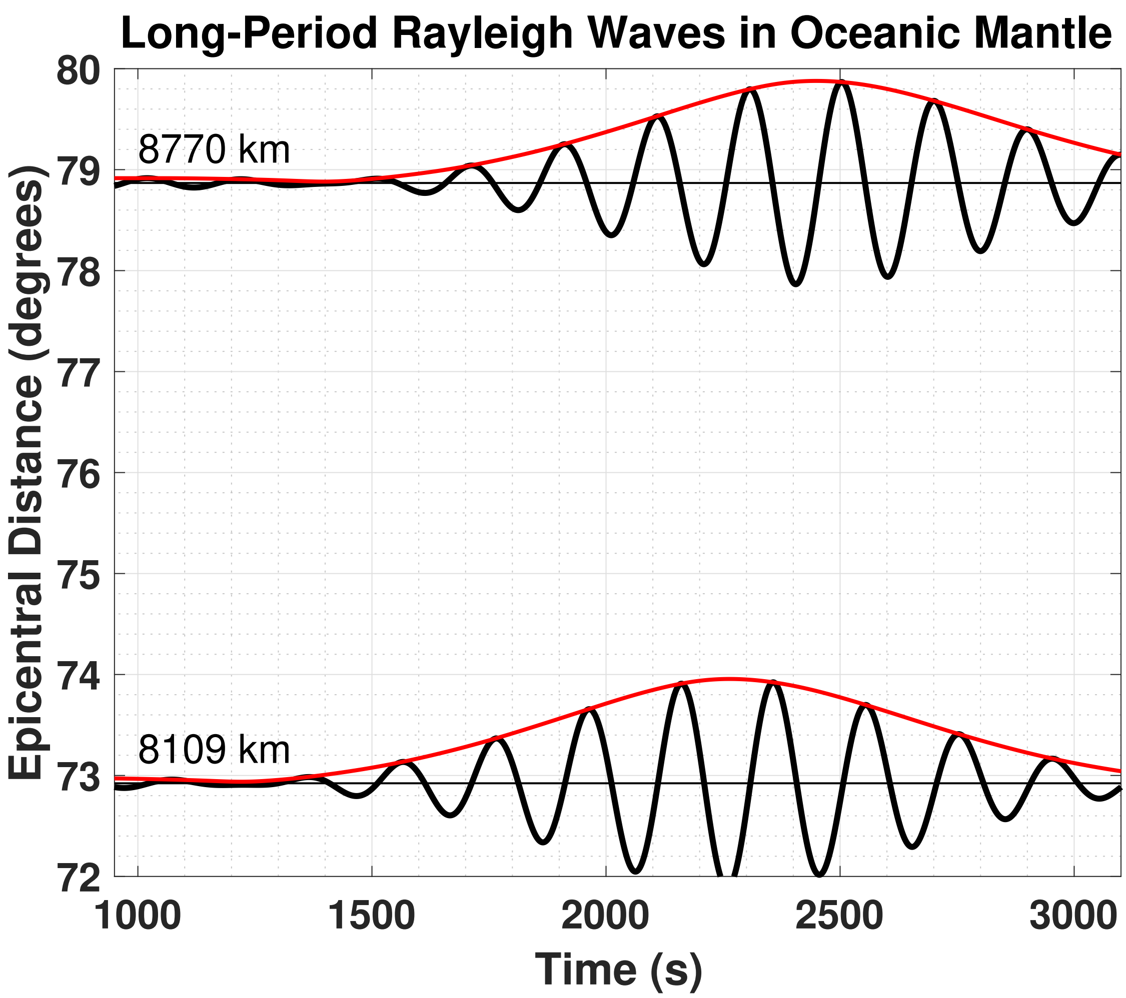
Section 1: Measuring Phase and Group Velocities from Seismograms

Before we begin actual tomography, you should have a sense of how one might generate the dataset you'll be working with, and what measuring a travel-time between two stations really means. The following seismograms show long-period Rayleigh waves at two stations (filtered around 200 s), propagating through oceanic mantle. The distance of each station relative to the earthquake is shown on the left, in km.

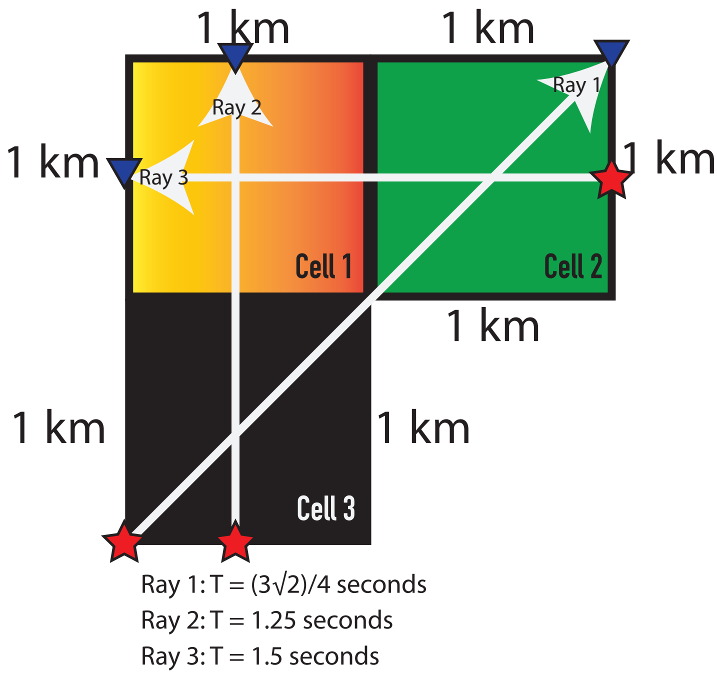
Q1: Measure the group and phase velocities in the mantle using this pair of stations.
Recall that:

- The **group** velocity corresponds to the speed of the **envelope that modulates the waveform**. To measure the group traveltime, measure the time that separates the maxima of the waveform envelope (shown in red) at each station.
- The **phase** velocity corresponds to the speed of the **individual peaks in the waveform**. To measure the phase traveltime, measure the time that separates the same individual peak of of the actual waveform (shown in black) at each station.

In both cases, you can divide the distance between the stations by the traveltime you have measured to obtain a phase or group velocity.



Section 2: Tomography by Hand



Consider the diagram above, which is to scale. The red stars are earthquakes and the blue triangles indicate seismic stations that record surface waves from the earthquakes.
Q2: Using traveltime measurements for the three rays shown, solve for the phase velocity in each cell.

You have the following information (also shown on the diagram).

Ray 1 passes through Cell 3 and Cell 2 and has a phase traveltime of $\frac{3}{4}\sqrt{2}$ **seconds.**

Ray 2 passes through Cell 3 and Cell 1 and has a phase traveltime of 1.25 **seconds.**

Ray 3 passes through Cell 2 and Cell 1 and has a phase traveltime of 1.5 **seconds.**

Feel free to solve for the velocities either by solving this problem as a conventional system of 3 equations with three unknowns, or by phrasing this as a matrix and taking the inverse of the matrix in MATLAB.

Section 3: Working with your Dataset of Rayleigh wave Measurements

About Your Dataset

Your dataset is derived from a study (Foster et al., 2014) that solved for a phase-velocity map over the US. Your measurements are **phase** traveltimes measured on Rayleigh waves at a period of 100 s, between pairs of USARRAY stations throughout the central Continental US.

Each measurement corresponds to a different pair of stations at distinct locations in the US that both record Rayleigh waves from the same Earthquake. Each pair of stations is separated by a provided, inter-station distance, and the traveltime of the Rayleigh wave between the two stations is measured and provided. Using this information, you will eventually solve for the phase-velocity map of the US which best explains these observations.

You have been provided with a .mat file containing this information. Try loading it into MATLAB using the following code.
Note: You may want to begin by making a new .m file using the 'new script' command on the top left of your MATLAB window, and storing all your following code inside this script so it can easily be repeated.

```
load('DataFile_1610Lab1.mat') % loads in your dataset as an array
```

This will generate an array in your workspace called 'Data'. Each row corresponds to a different measurement. This array has several columns;
(1) The longitude of the the first station (Station A) in a station pair for that measurement
(2) The latitude of the the first station (Station A) in a station pair for that measurement
(3) The longitude of the the second station (Station B) in a station pair for that measurement
(4) The latitude of the the second station (Station B) in a station pair for that measurement
(5) The inter-station distance between Station A and Station B (in km)
(6) The Rayleigh wave traveltime between each pair of stations (in seconds)
(7) The longitude of the midpoint between Station A and Station B
(8) The latitude of the midpoint between Station A and Station B
An example of how you can load this information into your workspace, as individual vectors, is shown below:

```
StationA_Lons = Data(:,1);
StationA_Lats = Data(:,2);
StationB_Lons = Data(:,3);
StationB_Lats = Data(:,4);
InterStnDists = Data(:,5);
Ttimes = Data(:,6);
MidPtLon = Data(:,7);
MidPtLat = Data(:,8);
```

(You will use the other columns in Section 4)

Tasks

Q3a: Make a crude map of station locations. Plot longitude on the x-axis and latitude on the y-axis. Plot Station A locations as blue squares and Station B locations as green stars.

To do this, you will need to use the MATLAB command **plot**. **Plot** takes in a list of x-coordinates followed by a list of y-coordinates. Any following arguments can change the appearance (such as color or marker type) of your plot. For example:

```
figure
plot(StationA_Lons,StationA_Lats,'bs')
hold on
plot(StationB_Lons,StationB_Lats,'g*')
grid on
xlabel('Longitude')
ylabel('Latitude')
```

```
legend('StationA','StationB')
```

This plots an x-y graph, with the variable x plotted on the x-axis and variable y plotted on the y-axis. The points are colored by the value in the third argument you specify. ('bs' results in blue squares and 'g*' results in green stars.)

Q3b: To your map in Q3a, add the station midpoints as little black dots. Comment on the distribution of midpoints versus the distribution of stations.

```
plot(MidPtLon,MidPtLat,'k.')
```

Q3c: Make a scatter plot showing the traveltime measurements as a function of inter-station distance. What are the shortest and longest station separations in your dataset? Why do you think the spread of traveltimes is small for the shortest station separations and large for the longest separations? Can you estimate a ballpark Rayleigh phase velocity from the values on this graph?
You can use the 'scatter' command to make your map. For example:

```
figure % this creates a new figure
scatter(InterStnDists,Ttimes,1,'filled')
grid on
xlabel('Station Separation (km)')
ylabel('Rayleigh Wave Travel Time (sec)')
```

Q3d: Calculate the speed traveled by the Rayleigh waves between every pair of stations. This is straightforward, since you have distance and time for each pair. Make a scatter plot showing speed as a function of inter-station distance. What are the low-end and high-end values for phase velocity? Do the values change with station separation, and if so, how? Speculate on reasons why. What is the disadvantage of interpreting phase velocity values for larger inter-station distances?

```
cavg=InterStnDists./Ttimes;
figure
scatter(InterStnDists,cavg,1,'filled')
grid on
xlabel('Station Separation (km)')
ylabel('Rayleigh Wave Speed (km/sec)')
```

Q3e: Explore whether there are any obvious trends in phase velocity as a function of longitude or latitude. Use the longitude and latitude of the station midpoints. Comment on any trends you observe.

```
figure
subplot(2,1,1)
scatter(MidPtLon,cavg,1,'filled')
grid on
xlabel('Station Midpoint Longitude')
ylabel('Rayleigh Wave Speed (km/sec)')

subplot(2,1,2)
scatter(MidPtLat,cavg,1,'filled')
grid on
xlabel('Station Midpoint Latitude')
ylabel('Rayleigh Wave Speed (km/sec)')
```

Q3f: Make a crude map of phase velocity as a function of longitude and latitude, using the station midpoints. Since you will be color-coding the points by phase velocity, the command "colorbar" will allow you to see the color scale and "caxis" will allow you to set limits on the color scale.

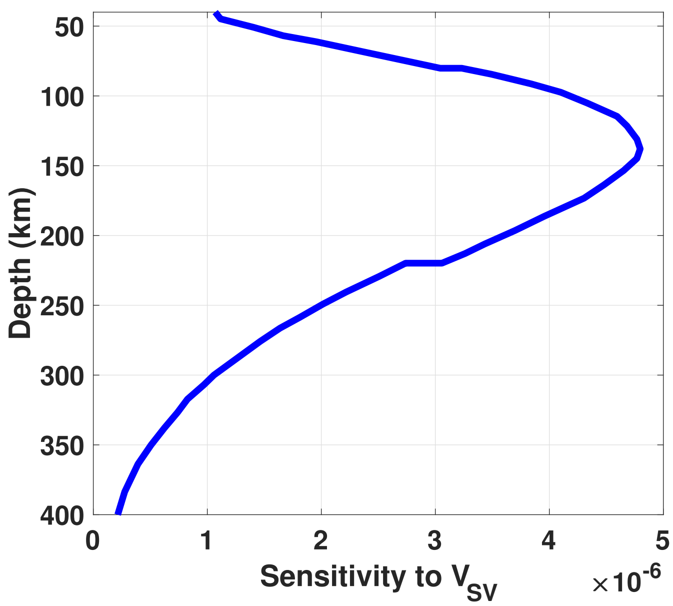
```
figure
scatter(MidPtLon,MidPtLat,10,cavg,'filled')
xlabel('Longitude')
```

```
ylabel('Latitude')
caxis([3.9 4.3])
colorbar
```

Q3g: Make the map a little less crude by plotting global coastlines and zooming into the study area. The 'xlim' and 'ylim' commands let you zoom into a specific axis range. The command 'load coast' loads a dataset consisting of the locations (longitudes and latitudes) of global coastlines into the variables 'long' and lat'.

```
load coast
hold on
plot(long,lat)
xlim([-105 -75])
ylim([25 50])
```

Q3h) The figure below is one that was shown to you previously in class, for Rayleigh waves at different periods. This version of the figure only shows the plot for a 100 s Rayleigh wave, which is most relevant to your measurements. This illustrates the depth range that the Rayleigh wave 'sees'. Using this figure, what depth (in km) is the Rayleigh wave most 'sensitive' to? In other words, what depth is your measurements most representative of? State a single value. To zeroth order, you may assume your measurements reflect shear velocity values at this depth.

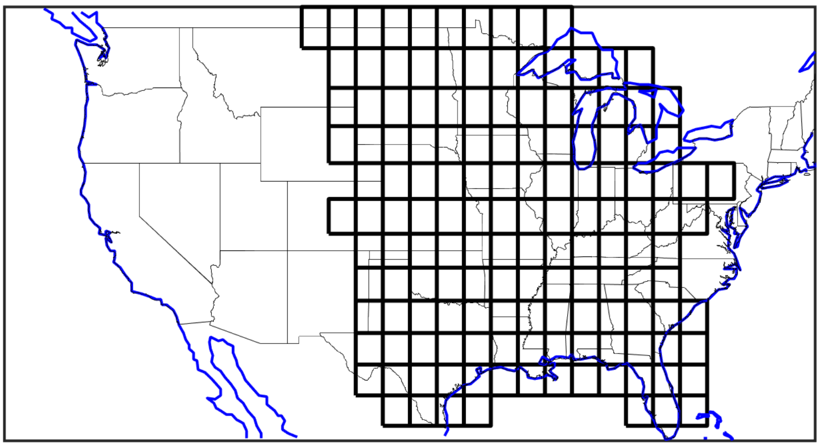


Q3i) Interpret your map of phase velocities from Q3g in terms of the tectonics of the continental US. How do velocities change from south to north, and what lithospheric feature do you think could be responsible?

If you need guidance with this, refer to the lecture06_092821.pdf file on canvas for a schematic map of the tectonics of North America from Schaeffer and Lebedev (2014).

Section 4 [Extra Credit, Optional]: Conducting Tomography on Real Rayleigh wave Measurements

The grid you will be working on is shown below. Using the dataset that you've worked with, you'll be solving for the phase velocity in each of these cells.



The file 'Grid.mat' contains the locations of the centers of each of these cells. To get this information, you'll need to load it into MATLAB, upon which it will create a variable TOMOGRID. The first column is the cell longitudes and the second column is the cell latitudes.

Q4a: Make your own map showing the centers of each of these grid cells, where each is plotted as a red triangle. How does the location of these cells compare with the locations of the stations from your plot in Q3a?

```
load('Grid.mat')
GridCenterLon = TOMOGRID(:,1);
GridCenterLat = TOMOGRID(:,2);
 figure()
 plot(GridCenterLon,GridCenterLat,'r^')
 hold on
 xlim([-105 -75])
 ylim([25 50])
```

Q4b: Add 5 example raypaths from any of your ≈ 360000 *measurements to your plot. The code below shows how you might do this for one of your measurements, corresponding arbitrarily to row 50000- the 50,000th measurement. Note that the only purpose of the '...' is to allow code to carry over to the next line and be readable. Can you make a rough guess as to approximately how many cells each raypath traverses, on average?*

```
measurement_num = 50000
plot([StationA_Lons(measurement_num) ...
    StationB_Lons(measurement_num)],[StationA_Lats(measurement_num) ...
    StationB_Lats(measurement_num)],'linewidth',4)
```

Q4c: Make a map showing the 'coverage' of your measurements, or the total amount of distance that all the measurements travel through each of your grid cells. What parts of your grid do your rays travel the most through? In general, grid cells that many surface waves pass through end up being more reliable in tomographic images than grid cells that don't have much ray coverage. Consequently, which section of your grid are you most confident in being able to accurately image?

To help you with this, we have pre-calculated the distance that each ray travels through each of these cells and stored it in the variable 'Data', for each row. Specifically, each column of the array 'Data' between columns 9 and 151 stores the distance that the ray corresponding to the measurement on that row travels through each of these cells.For example, the entry on column 9 on row 1 stores the distance that the ray on row 1 travels through Cell 1. The entry on column 10 on row 1 stores the distance that the ray on row 1 travels through Cell 2... and so on.

To isolate the information between columns 9 and 159 and store it in the variable 'L', you could do something like:

```
L = Data(:,9:159);
```

This generates a matrix that corresponds to the **L** matrix from the background section.

```
L = Data(:,9:159);
Coverage = sum(L);
figure()
scatter(GridCenterLon,GridCenterLat,450,Coverage,'filled','^')
colorbar
```

Q4d: Solve for phase velocity map using your traveltime measurements and a make a map of your results. Does this agree with your map from section 3? Are there any artifacts in your map? For instance, are there values that seem 'strange', perhaps anomalously large and not geologically reasonable? Speculate on what could cause this, and how this could be fixed. More generally, how do you think you could improve the quality of your image?

You now have the matrix that corresponds to the L. Remember that the traveltimes correspond to the **t** term, and that you want to solve for p, which is a vector of (1/velocity) values at each cell. You can solve any matrix equation like this quite simply in MATLAB: If $\mathbf{t} = \mathbf{Lp}$, then you solve for **p** using the following code: (after you have **p** you need to convert it into phase velocity by taking the reciprocal).

```
 p=L\Ttimes;
phvel = 1./p;
figure()
scatter(GridCenterLon,GridCenterLat,150,phvel,'filled','s')
colorbar
caxis([3.9 4.3])
```

Q4e: Do your new results agree with your map from section 3? Are there any artifacts in your map? For instance, are there values that seem 'strange', perhaps anomalously large and not geologically reasonable? Speculate on what could cause this, and how this could be fixed. More generally, how do you think you could improve the quality of your image?

Useful References

For more about the dataset that you used, refer to:

Anna Foster, Göran Ekström, Meredith Nettles, **Surface wave phase velocities of the Western United States from a two-station method**, Geophysical Journal International, Volume 196, Issue 2, February, 2014, Pages 1189–1206, https://doi.org/10.1093/gji/ggt454

For more about the theory of Seismic tomography, refer to:

Cliff Thurber & Jeroen Ritsema, 2007. **Theory and observations: seismic tomography and inverse methods**, Treatise on Geophysics, Vol. 1, pp. 323–360, Elsevier, Amsterdam.

For more about the most-up-to-date phase velocity maps of the continental US, refer to:

Jordyn C. Babikoff & Colleen. A. Dalton (2019). **Long-period Rayleigh wave phase velocity tomography using USArray.** Geochemistry, Geophysics, Geosystems, 20, 1990–2006. https://doi.org/10.1029/2018GC008073

Let Colleen or Anant know if you have trouble finding these, or any other papers, online, or if you have any questions related to the material covered in this in-class activity.