SIO 227B – Practical 3

This practical requires you to find the source mechanisms of some earthquakes. All the code is in prac3.dir in the class account and there are some GFS files in the directories 2000 and 2001. You should choose two events to do, either a deep and a shallow event or a smaller and larger event. Details of the events are given in the file "locale.cmt.big". This file has the origin time of the event (year, day, hr, min, sec) then the location (colatitude, longitude, depth) then the moment (in dyne-cm, divide by 10^7 to get Nm). The six elements of the moment tensor (as determined by the Harvard CMT project) follow along with a source duration and the last column is the scale factor you multiply the moment tensor elements by. It would be great if everyone could choose different events!

You should copy the events you are going to do into your own directory (use cp -r as the event files are really directories). You should use program "curse" to quickly run through the records and delete records that are bad, flag bad parts of records, or occasionally interpolate a spike. Please keep a record (for my benefit) of which stations you deleted and why). You should then run the script "filter.cmd" which takes the filename as an argument. This scripts low-passes the data and decimates to a 40sec sample interval (the originals are 10sec), then it high passes to remove low-frequency noise and chops out an appropriate 10 hour segment (see notes below). The script produces a file called "filv.gfs". You should edit this again with "curse" taking care to check that the first Rayleigh wave packet looks ok (if it doesn't then flag from the beginning of the record to the place where the data start to look good).

Once you are happy with your data file then run the script "filter2.cmd". This assumes the input file is called filv.gfs so if you have changed this, you will have to change the script. This program does a final low-pass of the data, then computes the greens functions (using program "green"), then filters the greens functions to match the data, then computes synthetics using the Harvard CMT solution. The script puts the greens functions in a file called "grn.in" and the synthetics in a file called "syn.in". You can now use program "mttc" to compute your own moment tensor solution (see below) and then use program "syndat" to generate your own synthetics.

You can view your synthetics using program recp which accepts the synthetic filename as an argument. You will almost certainly find that some stations are too noisy, some stations are flipped, some have incorrect instrument responses, some have timing errors. Often, the comparison with the synthetics shows the there is something wrong with the initial Rayleigh wave packet, Keep a record of what is wrong with each record. You can use program "iutil" to fix flipped records (in file filv.gfs). You should use "curse" to fix records with bad front ends (by flagging) and to delete the noisy records and the records with bad responses. You should then run the "filter2.cmd" script again and have another go at determining the moment tensor.

Here are some more details about the fitting procedure. Retrieving the moment tensor is a linear inverse problem which is over determined (i.e. you have more data than unknowns). The equation you will be solving is

$$a_j(\omega) = \sum_{i=1}^{6} B_{ij}(\omega)\psi_i(\omega)$$

which is equation 4.50 in the notes. a_j is the spectrum of the recording at the j'th station, B_{ij} is a matrix which we compute and the 6-vector, ψ, is the seismic moment-rate tensor. The many problems encountered in solving for the moment-rate tensor are discussed in the notes (page 98). Here, we consider some of the practicalities involved.

Low-frequency seismology requires fairly large earthquakes to get a usable signal-to-nose ratio (SNR). We typically use events with M_S greater than about 6.5 and there are usually about 20 such events per year. At very low frequencies (less than about 2mHz), ground noise rapidly increases as frequency decreases and there is only usable signal for extremely large earthquakes. Routine analysis of source mechanisms therefore uses frequencies above 2mHz. As frequency increases, the fundamental modes (which dominate the seismogram) sample more and more of the near-surface structure which also happens to be the most heterogeneous. Above about 6mHz, the error incurred by assuming the Earth is spherically symmetric becomes extremely large and the source is not accurately retrieved. We therefore fit the above equation in a 4mHz frequency band centered about 4mHz. The optimum record length for estimating the amplitude of a mode is about 0.5 Q cycles which is about 10 hours for a 250 second surface-wave equivalent mode.

When we construct Green's functions, we include all modes with frequencies less than 8mHz. This means that we can apply a low-pass filter with a corner at about 6mHz and get clean looking seismograms without the ringing asociated with a sharp spectral cut-off. The data must be filtered in exactly the same way so that a meaningful time-domain comparison can be made. The program that you use to filter the data is called NDEC. This program applies a zero phase shift convolution filter to your data. The Greens functions are computed using program GREEN and are then filtered with the same filter as you applied to the data in program FILGRN. Note that GREEN also applies the individual instrument responses to the Green's functions.

The linear fitting for the moment tensor elements is done in program MTTC. This program asks you several questions before using a singular value decomposition (SVD) to construct a generalised inverse of the matrix $\mathbf{B}$. One question is whether or not you want to normalise the data. This means that the average amplitude at each station is made to appear the same using row weights. This is sometimes a good idea because the fact that there is not much power at a station tells you as much about the source as when there is a lot of power. Least-square solutions tend to fit the large signals so that stations with little power will tend to be ignored in the fitting procedure unless the weighting is applied. Another question that is asked is whether or not you wish to apply a time-domain taper before FFTing. Tapering sharpens up the spectrum but this is one case where you don't necessarily want high resolution. The data, and the columns of the Greens functions are sums of decaying sinusoids so their spectra are spiky. Accurate source retrieval requires that the spikes be alligned in frequency. We have seen that 3D structure tends to move peaks around and anything which accentuates the misallignment of the Green's functions with the data is a bad thing – this includes tapering.

Now we get to the real fitting. The program will ask for a time constant (which is the baseline length of a triangle function – a reasonable shape for the moment-rate tensor elements) and solves the above equation. You can apply the deviatoric constraint so that no explosive or implosive component to the source is allowed and, if you have information about the orientation of a possible fault plane, you can apply a plane constraint. The fitting is done using a SVD which is described in several texts. Briefly, $\mathbf{B}$ is decomposed into left- and right-hand eigenvectors and a diagonal matrix of singular values, Λ, as

$$\mathbf{B} = \mathbf{U}\Lambda\mathbf{V}^T$$

where $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^T\mathbf{V} = \mathbf{I}$ where superscript T stands for transpose and $\mathbf{I}$ is the unit matrix. The generalized inverse is given by

$$\hat{\psi} = \mathbf{V}\Lambda^{-1}\mathbf{U}^T\mathbf{a}$$

which reduces to the standard least-squares solution when all singular values are taken. If there are some small singular values, the system of equations is not well-conditioned and we do not have enough information to completely constrain the solution. A reasonable response is to ignore all small singular values when constructing the solution. The fit to the data will probably not be degraded by much and the result is a solution with the smallest Euclidian norm for this fit to the data. The program asks for a "qmin-qmax" ratio which is the largest range of singular values you are prepared to accept. A value of 0.01 to 0.1 is appropriate and you can try experimenting with this parameter.

The program prints out some interesting numbers at each stage of the fitting, including fits of various recordings and an overall variance reduction. You should vary the time constant of the event until the overall variance reduction is optimised. Events with M_S below about 7 will probably have time constants of less than 30 seconds. The largest events can have time constants in excess of 100 seconds.

Once you have settled on a best fit, you can use program SYNDAT to compute synthetic seismograms. This program accepts double couples or moment tensors as input. You can view the fit of individual synthetics to the data using program RECP. This is where you may find some problems with the data. You may need to fine tune some editing or you may find that a time series has a timing error or an incorrect instrument response. The variance reductions printed out by MTTC for individual stations are useful for diagnosing data problems. If a variance ratio (the last column) is less than about 0.4, you will find that time-domain fits are excellent. A variance ratio between 0.4 and 0.7 is acceptable. 0.7 to 0.8 is barely acceptable and suggests a problem with the data. Above 0.8 means that there is almost certainly something wrong with the data or, if many recordings are badly fit, the event may have had a complex rupture history and cannot be modeled by a point source.

One last program for you to play with is DCSEARCH. This program was written to try to remove some of the bias associated with the linear fitting procedure. It takes a guess of a double couple model for the source (strike, dip and slip) and adjusts this until the power at each station is correctly predicted. This program calculates power in the time domain so you should use a relatively short time series including about two Rayleigh waves.

When writing this assignment up, please discuss how well-determined your solution is and how reasonable it is from a tectonic standpoint (remember the location of the event in file locale.cmt.big is given by colatitude and longitude). How close to a double-couple is it? Does you solution agree with the Harvard CNT solution. If your data is incapable of completely constraining the solution, try looking through the literature to get a better guess of source orientation – this might allow you to use the "plane constraint" to get a better idea of the moment. Discuss how well your solution fits the data and document any adjustments you made to your dataset during the fitting procedure.