PEAT8002 - SEISMOLOGY
Lecture 10: Earthquake relocation

Nick Rawlinson
Research School of Earth Sciences
Australian National University
Earthquake relocation is one of the classic inverse problems in seismology.

Until recently, most relocation procedures were based on exploiting arrival time information from the seismic signal produced by the earthquake.

Accurate earthquake relocation is vital for many applications, including hazard forecasting and assessment, seismic tomography, examination of the contemporary stress field, source mechanism studies etc.

In the past few years, it has been demonstrated that coda wave interferometry can be used for earthquake relocation. Using this approach, the shape of the seismic waveform is exploited rather than the onset times of specific phases.
If we approximate an earthquake as a point source, then its location is defined by four parameters - three spatial coordinates and an origin time. In Cartesian coordinates, the hypocenter location is \((x, y, z, t_0)\), while in spherical coordinates, it is \((r, \theta, \phi, t_0)\).

The epicenter is the location \((x, y)\) or \((\theta, \phi)\) on the Earth’s surface above the hypocenter.

Most earthquakes occur within about 30 km of the Earth’s surface. Deep focus earthquakes occur down to about 700 km depth, but no seismic events have been recorded from depths greater than this.

An initial estimate of distance from a recorder and origin time can be determined from the difference in arrival time between P and S phases.
In fact, it is possible to exploit the differential moveout effect that the spherically symmetric Earth has on any identifiable phase in a seismogram.
The difference in arrival times between the different phases can be matched to standard traveltime tables.

This will provide an estimate of origin time and the angular distance between source and receiver.

Repeating this procedure for multiple stations allows a location estimate to be made from the multiple intersecting circles of constant angular distance.
The above procedure yields an approximate location and origin time for an earthquake recorded by a regional or global network of stations. However, greater accuracy is often required in many applications.

In addition, at smaller distances, it can be difficult to associate later arrivals with known phases, and the underlying velocity model is often poorly known.

By formulating the earthquake relocation problem as an inverse problem, it is possible to obtain much more accurate estimates of hypocenter location.

The basis of the inverse problem is the following relationship between data $d$ and model parameters $m$:

$$d = g(m)$$
In this case, the data matrix \( d \) are the arrival times \( T_i \) \((i = 1, \ldots, N)\) of one or more phases generated by the earthquake at an array of stations.

The model parameter vector \( m \) is simply \( m = (r, \theta, \phi, t_0) \), the hypocenter of the earthquake.

For a given velocity model, adjusting the location and/or origin time of an earthquake will obviously modify \( d \), so clearly \( d = g(m) \).

In cases where the velocity model is poorly known in addition to hypocenter coordinates, it is often necessary to simultaneously invert for both parameter classes. However, in the following treatment, we will ignore this possibility.
Put simply, the inverse problem is to adjust the model parameters $\mathbf{m}$ until the data $\mathbf{d}$ are satisfied.

This presumes that we have a method for predicting arrival times given some arbitrary value for $\mathbf{m}$.

In the case of teleseismic events, this could be the standard traveltime tables associated with global reference models such as ak135 or iasp91.

Generally, it is preferable to use the most accurate model available. If this includes lateral heterogeneity, then ray or wavefront tracking methods (e.g. FMM) can be used to predict traveltimes.

For absolute hypocenter locations, standard approaches include iterative non-linear relocation and fully non-linear relocation.
Given some initial estimate of an earthquake hypocenter $m_0$, the general relationship $d = g(m)$ can be written as a Taylor series expansion:

$$d = g(m_0) + G(m - m_0) + O(m - m_0)^2$$

where $G = \partial g / \partial m$ is a matrix of partial derivatives.

If we define $\delta d = d - g(m_0)$ and $\delta m = m - m_0$, then to first order:

$$\delta d = G \delta m$$

This new equation assumes that a linear relationship exists between the arrival time data and the hypocenter locations.
Provided that the matrix of partial derivatives is known, it is possible to solve the linear system of equations. However, in general $G$ is not square because usually there are more data than unknowns. For $N >> 4$ traveltime values,

$$
\begin{bmatrix}
\delta d_1 \\
\delta d_2 \\
. \\
. \\
\delta d_N
\end{bmatrix} =
\begin{bmatrix}
G_{11} & G_{12} & G_{13} & G_{14} \\
G_{21} & G_{22} & G_{23} & G_{24} \\
. & . & . & . \\
. & . & . & . \\
G_{N1} & G_{N2} & G_{N3} & G_{N4}
\end{bmatrix}
\begin{bmatrix}
\delta m_1 \\
\delta m_2 \\
\delta m_3 \\
\delta m_4
\end{bmatrix}
$$

The above system of equations has many more data than unknowns (assuming linear independence), and therefore constitutes an overdetermined inverse problem.
One way to overcome this problem is to seek the least squares solution to the above system. Thus, we seek the model adjustment that gives the best match (in the least squares sense) to the set of observations (cf. linear regression).

If we define a function $S(\delta m)$ as:

$$S(\delta m) = (\delta d - G\delta m)^T(\delta d - G\delta m) = \|\delta d - G\delta m\|^2$$

then the least squares solution corresponds to the model perturbation $\delta m$ that minimises $S(\delta m)$. As in linear regression, this corresponds to when $\nabla S(\delta m) = 0$. 

Earthquake relocation
Iterative non-linear relocation
The least squares function $S(\delta m)$ can be written in index notation as follows:

$$S(\delta m) = \sum_{j=1}^{N} \left[ \delta d_j - \sum_{i=1}^{4} G_{ji} \delta m_i \right]^2$$

The relationship between the notation and matrix forms is easy to see from:

$$S(\delta m) = \begin{bmatrix} \cdots & \delta d_j - \sum_{i=1}^{4} G_{ji} \delta m_i & \cdots \end{bmatrix} \begin{bmatrix} \delta d_j - \sum_{i=1}^{4} G_{ji} \delta m_i \\ \cdots \\ \cdots \end{bmatrix}$$

as each matrix has $N$ elements.
If we now differentiate with respect to a particular model parameter $\delta m_k$, then we can use the chain rule to obtain the derivative:

$$\frac{\partial S(\delta m)}{\partial \delta m_k} = \frac{\partial S(\delta m)}{\partial u_j} \frac{\partial u_j}{\partial m_k}$$

where we have made the substitution:

$$u_j = \delta d_j - \sum_{i=1}^{4} G_{ji} \delta m_i$$

For the first term we obtain:

$$\frac{\partial S(\delta m)}{\partial u_j} = 2 \sum_{j=1}^{N} \left[ \delta d_j - \sum_{i=1}^{4} G_{ji} \delta m_i \right]$$
Earthquake relocation
Iterative non-linear relocation

- The second term yields:
  \[ \frac{\partial u_j}{\partial m_k} = G_{jk} \]
  since the derivative of \( G_{ji} \delta m_i \) is non-zero only when \( i = k \).
- Putting these two terms together and setting to zero yields:
  \[ 2 \sum_{j=1}^{N} \left[ \delta d_j - \sum_{i=1}^{4} G_{ji} \delta m_i \right] G_{jk} = 0 \]
  This may be re-written as:
  \[ \sum_{j=1}^{N} \sum_{i=1}^{4} G_{ji} G_{jk} \delta m_i = \sum_{j=1}^{N} G_{jk} \delta d_j \]
Given that the above expression is valid for \( k = 1, \ldots, 4 \), it may be written in matrix notation as:

\[
G^T G \delta m = G^T \delta d
\]

The least squares solution to the linearised inverse problem can therefore be written as:

\[
\delta m = [G^T G]^{-1} G^T \delta d = G^{-g} \delta d
\]

where the operator \( G^{-g} \) is often referred to as the \textit{generalised inverse}.

The model perturbation \( \delta m \) can be obtained using standard solvers such as conjugate gradients or SVD.
Earthquake relocation
Iterative non-linear relocation

- For some given initial model \( m_0 \), theoretical arrival times are calculated and compared to the observed times to produce \( \delta d \). The matrix \( G \) is also evaluated at this stage.
- This allows the model perturbation \( \delta m \) to be produced by solving the above equation. The new model \( m_1 = m_0 + \delta m \).
- Due to the inherent non-linearity of the inverse problem, the above procedure must in general be iterated until it converges on a solution. Thus, theoretical arrival times must be computed for the new model \( m_1 \) and the matrix \( G \) updated prior to obtaining \( m_2 \).
- The model update procedure is defined by:

\[
m^{i+1} = m^i + \delta m^i
\]
Earthquake relocation
Iterative non-linear relocation

- Note that because we have assumed local linearity to obtain a solution at each iteration, the function $S(m)$ must be smooth and well behaved (quasi-quadratic).
- In practical applications, this iterative non-linear approach is usually only successful when a good initial estimate of hypocenter location is available, and/or the velocity medium is not strongly heterogeneous.
- In the presence of strong velocity gradients and data noise, $S(m)$ is often characterised by multiple minima, which makes solution difficult.
- Nevertheless, iterative non-linear location is still frequently used in practice.
In 3-D space, we need to derive expressions for four partial derivatives in order to construct $G$.

The derivative of arrival time $T$ with respect to origin time $t_0$ is given by:

$$\frac{\partial T}{t_0} = 1$$

since perturbing the origin time will perturb the arrival time by the same amount. In other words, $T = t_o + t_{tr}$ where $t_{tr}$ is source receiver traveltime which is independent of $t_0$. Clearly, $\partial(t_o + t_{tr})/\partial t_0 = 1$.

The remaining terms can be computed analytically to first order. In this case, our derivation will assume a Cartesian coordinate system.
Earthquake relocation
Calculation of $\mathbf{G}$

Consider a perturbation in hypocenter depth given by $\Delta z_h$. To first order, the corresponding change in path length $\Delta l$ is given by:

$$\Delta l = -\Delta z_h \cos \theta_h$$

where $\theta_h$ is the angle between the ray and the $z$ – axis at the hypocenter.

Note that the change in path length is negative as the hypocenter is perturbed to shallower depth.
The change in arrival time $T$ is therefore given by:

$$\Delta T = \frac{\Delta l}{v_h} = -\frac{\Delta z_h \cos \theta_h}{v_h}$$

where $v_h$ is the velocity at the hypocenter.

The first-order accurate expression for the partial derivative is therefore given by:

$$\frac{\partial T}{\partial z_h} = -\frac{\cos \theta_h}{v_h}$$
Similarly, the partial derivatives for the two remaining parameters are:

\[
\frac{\partial T}{\partial x_h} = -\frac{\cos \phi_h}{v_h}, \quad \frac{\partial T}{\partial y_h} = -\frac{\cos \psi_h}{v_h}
\]

where \( \phi_h \) and \( \psi_h \) subtend the horizontal projection of the ray path and the \( x \) and \( y \) axes, respectively, at the hypocenter.

Most local earthquake tomography schemes that invert for both velocity and hypocenter location use this form of partial derivative.

For regional and global scale relocation, the derivatives should be formulated in spherical coordinates.
In the synthetic example below, five hypocenters are relocated using the above procedure in the presence of significant lateral heterogeneity. FMM is used to predict arrival times in this case.
In most realistic media, the relationship between data and model parameters $d = g(m)$ is highly non-linear, so if the initial location is poor, iterative non-linear methods are likely to fail (i.e. not converge or converge to the wrong minimum).

An alternative means of hypocenter location is to search model space without relying on local gradient information.

One simple way of doing this is to perform a basic grid search. Given the speed of modern computers, this can be a practical approach, particularly when coupled with grid based solvers like FMM.

Otherwise, one can use non-linear/global optimisation techniques such as genetic algorithms, simulated annealing and the neighbourhood algorithm.
Consider an objective function $S(m)$ defined by:

$$S(m) = (d_{\text{obs}}^r - d_m^r)^T (d_{\text{obs}}^r - d_m^r)$$

where $d_{\text{obs}}^r$ is a vector of observed arrival times (from a given hypocenter) at a set of $N$ stations with the mean subtracted, and $d_m^r$ is a vector of predicted travel times, also with the mean subtracted.

Removing the mean is useful, because it eliminates the origin time as an unknown in the inversion. This is because $d_{\text{obs}} = d^t + d_0$ where $d_0$ is the origin time and $d^t$ is the observed traveltime.

Thus, the mean of $d_{\text{obs}}$ is

$$\bar{d}_{\text{obs}} = \frac{\sum_{i=1}^{N} d_i^t}{N} + d_0$$
Similarly, the mean of \( \mathbf{d}_m \) is

\[
\bar{\mathbf{d}}_m = \frac{\sum_{i=1}^{N} d^i_m}{N}
\]

In this case, there is no \( d_0 \) term because \( \mathbf{d}_m \) is the vector of predicted traveltimes.

Thus, from the above expressions:

\[
\mathbf{d}^r_{\text{obs}} = \mathbf{d}^t - \bar{\mathbf{d}}^t, \quad \mathbf{d}^r_m = \mathbf{d}_m - \bar{\mathbf{d}}_m
\]

Thus, the objective function defined above only compares traveltime information. Once \( S(\mathbf{m}) \) has been minimised, the origin time can be found simply from \( \mathbf{d}_0 = \mathbf{d}_{\text{obs}} - \mathbf{d}^t \).
The origin time does not affect the relative arrival time pattern, and therefore no information is ignored by using this approach.

In the following 2-D examples, FMM is used to compute traveltimes in the presence of velocity models of varying complexity.

Traveltimes from a regular grid of points (180,901 in total) that span the model are computed to every receiver that lies on the surface (5 in this case). Rather than execute FMM 180,901 times, we can use each receiver as a virtual source to compute all required traveltimes.

In each example, we try to find the location of the earthquake, which lies in the centre of the model, using only the arrival times at five receivers lying on the surface.
The first example uses a simple model in which velocity varies linearly with depth.
Model traveltime predictions for every receiver-grid point pair are obtained by appealing to the reciprocity principle.
Earthquake relocation

Example 1

Contours of the objective function are illustrated below; the minimum clearly corresponds with the true source location.
In the second example, a model that features very strong lateral variations in wavespeed is used.
Contours of the objective function are more complex in this case, but a clear minimum is still present.
In the final example, a pathological velocity model is used, which has velocity contrasts as large as 70:1.
Contours of the objective function have a strong relationship with the velocity model, but nonetheless, the correct minimum zone can be identified.
The above examples demonstrate the potential of both FMM and fully non-linear location methods in earthquake location.

A robust non-linear optimisation scheme should be capable of locating the minimum of the objective function even in the presence of strong velocity gradients.

Poor station coverage, significant velocity heterogeneity, and data noise may result in multiple minima; techniques like the neighbourhood algorithm explore model space and produce ensembles of solutions, which can then be interrogated for their robustness.
Relative earthquake location techniques are most usefully applied in seismogenic zones, such as in the neighbourhood of convergent or strike-slip plate margins. Although they can only find the relative locations of multiple earthquakes, they can do so with much greater accuracy than absolute location techniques.

Most of these techniques rely on the assumption that if the hypocentral separation between two earthquakes is small compared to the distance to the receiver, travel time differences at a particular station are not biased by the effects of heterogeneity.

In other words, ray trajectories are almost identical, with the major difference simply being the change in path length. Thus, the difference in travel times for the two events can be attributed to their offset.
Waveform cross-correlation techniques can also be used to improve the accuracy of relative arrival time picks.

The assumption of coherent waveforms for two different events is often most valid when they are very close together; source mechanisms can be identical and scattering effects are similar.

One popular relative relocation scheme is the so-called double-differencing technique. It exploits P- and S-wave differential traveltime measurements.

Residuals between observed and theoretical traveltime differences (or double-differences) are minimised for pairs of earthquakes at each station, while linking together all observed event-station pairs.
Figure 9. (a) NCSN locations and (b) double-difference locations obtained from catalog travel-time differences and (c) the combined set of catalog and cross-correlation data of events located in the El Cerrito and the Berkeley cluster. Top panels show map view, lower panels show cross sections along the fault in NW–SE direction. The surface trace of the Hayward fault and shore line is shown. Boxes indicate events included in the cross sections in this Figure and in Figure 11.