

Notizen zur Vorlesung

Seismische Tomographie

gehalten von
Dr. Florian Bleibinhaus
LMU München
SS 2004

15. Juli 2004

geschrieben von
Christoph Moder

<http://www.skriptweb.de>

Hinweise (z.B. auf Fehler) bitte per eMail an uns: mail@skriptweb.de – Vielen Dank.

1 Was ist seismische Tomographie?

- Tomographische Verfahren sind aus der Medizin bekannt, z.B. die Computertomographie (Röntgenstrahlen) oder die Kernspintomographie (Magnetfelder). Davon inspiriert werden tomographische Methoden auch in der Seismik angewandt, unter Verwendung von seismischen Wellen.
- Schwierigkeiten: In der Medizin kann die zu untersuchende Person von allen Seiten durchleuchtet werden, während man bei der Seismik nur von einer Seite, nämlich der Erdoberfläche, Zugang hat. Außerdem können die medizinischen Geräte den Körper scheibenweise durchleuchten (d.h. zweidimensional), während man es in der Seismik immer mit 3D-Durchleuchtungen zu tun hat.
- Prinzip der Tomographie: Man durchleuchtet das Untersuchungsobjekt und misst dabei die Signalamplitude, die auf der anderen Seite ankommt. Dadurch weiß man zunächst nur die über den Laufweg aufsummierte Dämpfung, aber nichts über die Verteilung der dämpfenden Komponenten (vgl. Röntgenbild: man erhält eine 2D-Projektion ohne Höheninformationen). Zunächst nimmt man daher eine gleichmäßig über den Laufweg verteilte Dämpfung an. Da man das Untersuchungsobjekt aber in verschiedenen Richtungen durchleuchtet, wird jeder Punkt von mehreren Strahlen durchleuchtet, und daraus lässt sich die dritte Dimension errechnen. (Die Kernspintomographie unterscheidet sich in dieser Hinsicht etwas, weil sie direkt Tiefeninformationen liefert – weil die Feldstärke des angelegten Magnetfelds nicht konstant ist, sondern einen Gradienten hat, und die Frequenz des Kernspins von der Feldstärke abhängt, wodurch man aus der gemessenen Frequenz auf die Position im Magnetfeld schließen kann. Mit dieser Zusatzinformation sind schärfere Bilder möglich.)
- Während man z.B. in der Röntgentomographie die Amplitude des Röntgenstrahls, der das Untersuchungsobjekt durchquert hat, misst, ist es bei der seismischen Tomographie die Laufzeit (konkret: die Zeit vom Schuss bis zum Ersteinsatz auf dem Seismogramm), die betrachtet wird. Der Grund dafür ist erstens, dass die Geschwindigkeit seismischer Wellen vor allem von der Dichte des Gesteins abhängt, während ihre Amplitude auch von anderen Faktoren abhängt und darum über die Dichte keine so eindeutigen Aussagen macht. Zweitens hat man bei der Röntgentomographie gerade Laufwege (die Krümmung des Strahls durch verschiedene optische Dichten im Material ist so winzig, dass man sie vernachlässigen kann), während in der Seismik die Strahlen sehr deutlich gekrümmt sind und auch gestreut oder fokussiert werden, was die Interpretation der Amplitude deutlich erschwert.
- Oft modelliert man nicht nur die Geschwindigkeit der seismischen Wellen, sondern gleichzeitig die Gesteinsdichte und die Signalamplitude, so dass man das seismische Modell anhand der gemessenen Amplitude und Gravimetrie-Messdaten überprüfen kann.
- Die Tomographie ist ein sogenanntes „Rückwärtsproblem“ (back projection): Man versucht, aus der Beobachtung ein mögliches Modell herzuleiten (statt umgekehrt mit einem vorhandenen Modell die Beobachtungen vorherzusagen).
- Mathematisch gesehen entspricht eine tomographische Durchleuchtung einer Radon-Transformation (J. Radon, 1917):

$$R f(r, \theta) = \int_S^R f(r, \theta') dz$$

Die Betrachtung erfolgt in Polarkoordinaten, r ist die Entfernung zum Ursprung, θ der Drehwinkel.

- Man misst also die Radon-Transformation der Funktion $f(r, \theta)$ und möchte daraus die ursprüngliche Funktion rekonstruieren. Die Dichteverteilung, die man wissen möchte, befindet sich im Ortsraum (x, y) , das Ergebnis der Radon-Transformation befindet sich jedoch im Phasenraum (x, t) (d.h. das Seismogramm ist im Phasenraum). Durch eine 1D-Fourier-Transformation kann man es in den (k, ω) -Raum transformieren, und von dort aus über eine 2D-Fouriertransformation zurück in den Ortsraum.
Gleichung für die 1D-Fouriertransformation:

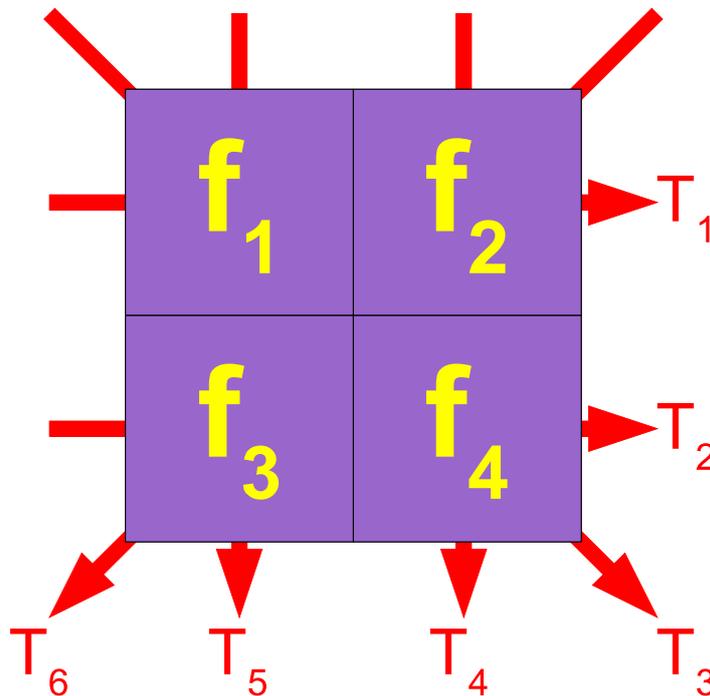
$$T'(k, \theta) = \iint f(x, y) e^{-2i\pi k(x \cos \theta + y \sin \theta)} dx dy$$

Üblicherweise wird eine Substitution durchgeführt: $u := k \cdot \cos \theta$, $v := k \cdot \sin \theta$

filtered back projection?

2 Tomographie an einem Beispiel

Ausgangssituation: Ein Körper mit quadratischem Querschnitt, der in vier Bereiche unterteilt ist, deren Absorptionskoeffizienten für die Durchleuchtungsstrahlen f_1, f_2, f_3, f_4 betragen. Dieser Körper wird aus sechs Richtungen durchleuchtet – zweimal horizontal, zweimal vertikal und zweimal diagonal; die gemessenen Absorptionen lauten $T_{1...6}$.



Nun definiert man einen sogenannten *sensitivity kernel*, der die einzelnen Absorptionskoeffizienten mit der gesamten Absorption pro Laufweg verknüpft. Das bedeutet, dass im sensitivity kernel z.B. Informationen über die Länge des Laufwegs im jeweiligen Bereich oder Richtungsinformationen (z.B. wenn das Absorptionsverhalten unterschiedlich ist, je nachdem in welcher Richtung man durchstrahlt) gespeichert sind:

$$s_k^{ij} = \frac{\partial T_{ij}}{\partial f_k}$$

Bedeutung der Indices: s_k^{ij} speichert Informationen darüber, wenn der Bereich k von i nach j durchquert wird. Es ergibt sich für die gemessene Absorption (Sender S_i , Empfänger R_j):

$$T(S_i, R_j) = \sum_{k=1}^n f_k s_k^{ij}$$

Man summiert also für jeden durchquerten Bereich das Produkt des dortigen Absorptionskoeffizienten mit dem jeweiligen Element des sensitivity kernels auf.

Weil man in der seismischen Tomographie mit Laufzeiten und Geschwindigkeiten rechnet, haben die Variablen folgende konkrete Bedeutungen: Die Messgröße T ist die Laufzeit der seismischen Wellen, der sensitivity kernel s enthält die Länge des Laufwegs, und der Absorptionskoeffizient f ist hier die sogenannte *Slowness*, das ist die reziproke Geschwindigkeit $u = 1/v$. Dann sagt die Gleichung aus: Laufzeit ist Laufweg mal Slowness.

Man kann den Zusammenhang auch vektoriell darstellen:

$$\vec{T} = \overleftarrow{s} \vec{f}$$

Ausgeschrieben bedeutet das:

$$\underbrace{\begin{pmatrix} 4,5 \\ 9,5 \\ 6,5 \cdot \sqrt{2} \\ 7 \\ 7 \\ 7,5 \cdot \sqrt{2} \end{pmatrix}}_{\vec{T}} = \underbrace{\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ \sqrt{2} & 0 & 0 & \sqrt{2} \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & \sqrt{2} & \sqrt{2} & 0 \end{pmatrix}}_{\overleftarrow{s}} \underbrace{\begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix}}_{\vec{f}}$$

Die Werte der Laufzeit \vec{T} wurden in diesem Beispiel beliebig gewählt. Die Matrix s speichert die Weglängen; die erste Zeile bedeutet beispielsweise, dass der f_1 -Bereich in der Länge 1 durchquert wird, ebenso der f_2 -Bereich. f_3 und f_4 werden vom T_1 -Strahl nicht durchquert, darum sind sie auf 0 gesetzt. Die dritte Zeile der Matrix beschreibt den T_3 -Strahl, der diagonal verläuft, darum sind die entsprechenden Weglängen auch $\sqrt{2}$.

Wenn man nun dieses lineare Gleichungssystem löst (z.B. indem man eine erweiterte Matrix schreibt und es mit dem Gaußschen Eliminationsverfahren auf Zeilenstufenform bringt), ergeben sich beispielsweise die Slowness-Werte $f_1 = 2, f_2 = 2,5, f_3 = 5, f_4 = 4,5$. Diese Werte sind linear abhängig, d.h. man braucht weitere Durchleuchtungsstrahlen (bzw. Strahlen auf anderen Wegen), um eine eindeutige Lösung zu erhalten.

3 Berücksichtigung von Fehlern

Wenn die beteiligten Werte fehlerbehaftet sind, liefert das lineare Gleichungssystem keine Lösung mehr. Fehlerhafte Werte sind jedoch der Normalfall, erstens weil Messfehler immer auftreten, und zweitens weil das gewählte Modell (z.B. im obigen Fall die Unterteilung in vier Bereiche) die Wirklichkeit nicht exakt darstellt. Als Ausweg addiert man noch einen Fehlervektor hinzu:

$$\vec{T} = \overleftarrow{s} \vec{f} + \vec{\varepsilon}$$

Anschließend versucht man, das Residuum ε zu minimieren:

$$\varepsilon^2 = \left| \overleftarrow{s} \vec{f} - \vec{T} \right|^2$$

Dieses Verfahren entspricht dem linearen Fit in x - y -Diagrammen, wo man ebenfalls die Fehler, nämlich die Abstände zwischen den gemessenen Werten und den berechneten Werten (d.h. die Punkte auf der Regressionsgerade mit gleichen x -Werten wie die Messwerte), zu minimieren versucht.

4 Nichtlinearität

Die Matrix $s_{lk} = s_{lk}(\vec{f})$ ist eine nichtlineare Funktion, d.h. eine Linearkombination ist nicht möglich:

$$s_{lk}(\vec{f}) + s_{lk}(\vec{g}) \neq s_{lk}(\vec{f} + \vec{g})$$

Man kann also die Beträge der Laufwege nicht einfach addieren, weil die Laufwege nach dem Fermatschen Prinzip bei verschiedenen Slowness-Werten unterschiedlich sind.

Laufzeit, ausgedrückt durch die Slowness $u = 1/v$:

$$t = \int_{\Gamma(\vec{v})} \frac{d\vec{r}}{\vec{v}(\vec{r})} = \int_{\Gamma(\vec{u})} \vec{u}(\vec{r}) d\vec{r}$$

Infinitesimal: $\vec{u} = \vec{u}_0 + \delta\vec{u}$; Taylor-Entwicklung um \vec{u}_0 :

$$t = \int_{\Gamma(\vec{u}_0 + \delta\vec{u})} \vec{u}_0(\vec{r}) d\vec{r} + \int_{\Gamma(\vec{u}_0 + \delta\vec{u})} \delta\vec{u}(\vec{r}) d\vec{r}$$

$$t = \int_{\Gamma(\vec{u}_0)} \vec{u}_0(\vec{r}) d\vec{r} + \int_{\Gamma(\vec{u}_0)} \delta\vec{u}(\vec{r}) d\vec{r} + \underbrace{\frac{\partial}{\partial \Gamma} \left(\int_{\Gamma(\vec{u}_0)} \vec{u}_0(\vec{r}) d\vec{r} \right)}_{:=0} \delta\Gamma + \underbrace{\frac{\partial}{\partial \Gamma} \left(\int_{\Gamma(\vec{u}_0)} \delta\vec{u}(\vec{r}) d\vec{r} \right)}_{:=0} \delta\Gamma + O^2$$

Dies ist die Taylor-Entwicklung bis zur 1. Ordnung. Terme höherer Ordnung vernachlässigt man, ebenso den vierten Summanden (weil dort eine Ableitung und etwas infinitesimal Kleines vorhanden ist). Außerdem gilt das Fermatsche Prinzip, das besagt, dass die Laufzeit von Strahlen extremal ist, d.h. ein Minimum oder auch Maximum annimmt – die Ableitung der Laufzeit muss also Null sein. Darum setzt man den dritten Term ebenfalls gleich Null.

Erläuterung zum Fermatschen Prinzip: Angenommen, man hat im Boden einen elliptischen Reflektor, und sowohl Schusspunkt als auch Geophonpunkt liegen in jeweils einem Brennpunkt. Dann sind alle Laufwege zwischen Schusspunkt und Geophonpunkt, die über den Reflektor gehen, gleich lang, ebenso die Laufzeit. Sie hat kein Maximum oder Minimum, der Reflektor bildet die Isochrone. Wenn man nun den Reflexionspunkt eines Strahls am Reflektor betrachtet und den Reflektor geringer krümmt, so wird der Strahl nach wie vor am selben Punkt reflektiert (weil sich dieser Punkt des Reflektors als einziger nicht geändert hat), für die Nachbarstrahlen hat sich jedoch der Laufweg vergrößert, da links und rechts des betrachteten Reflexionspunkts der Reflektor weiter entfernt als vorher ist. Der betrachtete Reflexionspunkt stellt nun ein Laufzeitminimum dar. Entsprechend ist es, wenn statt dessen der Reflektor enger gekrümmt wird – die Nachbarstrahlen haben nun einen kürzeren Weg bis zum Reflektor (und damit kürzeren Gesamtweg), damit hat der betrachtete Reflexionspunkt ein Laufzeitmaximum.

5 Dämpfung

$$\Delta \vec{T} = s_{lk} \Delta \vec{f}$$

Die Transponierte des sensitivity kernels links dranzumultiplizieren:

$$s_{kl}^T \Delta \vec{T} = s_{kl}^T s_{lk} \Delta \vec{f}$$

Nach f auflösen:

$$\Delta \vec{f} = (s_{kl}^T s_{lk})^{-1} s_{kl}^T \Delta \vec{T}$$

Der Ausdruck in der Klammer ist die *generalisierte Inverse*.

$$s_{lk} := U \Lambda V^T$$

$$\Delta \vec{f} = V \Lambda^{-1} U^T \Delta \vec{T}$$

Λ ist eine diagonale Matrix aus Eigenwerten, λ_n ist der Eigenwert aus $s_{kl}^T s_{lk}$

$$\Lambda^{-1} = \begin{pmatrix} 1/\lambda_1 & & 0 \\ & 1/\lambda_2 & \\ 0 & & \ddots \end{pmatrix}$$

Hinzufügen eines Dämpfungsterms θ , damit Λ bei kleinen λ nicht gegen ∞ , sondern gegen 0 geht:

$$\varepsilon^2 = \left| s\vec{f} - \vec{T} \right|^2 + \theta^2 \left| \Delta\vec{f} \right|^2$$

Damit wird aus:

$$s_{kl}^T \Delta\vec{T} = s_{kl}^T s_{lk} \Delta\vec{f}$$

folgender Term:

$$s_{kl}^T \Delta\vec{T} = \left(s_{kl}^T s_{lk} + \theta^2 \mathbf{1} \right) \Delta\vec{f}$$

6 Auflösungsmatrix und irreguläre Gitter

Im obigen Beispiel wurde das Quadrat durch ein rechtwinkliges Gitter in vier Bereiche unterteilt. Auch in der Realität verwendet man meist rechtwinklige Gitter, bei denen jedoch die Abstände zwischen den Zeilen bzw. Spalten nicht gleich sein müssen. Die Kreuzungspunkte der Gitterlinien geben die Geschwindigkeiten an, zwischen ihnen wird interpoliert (meist trilinear, aber auch Splines finden Verwendung, wenn Stetigkeit bei der Geschwindigkeitsveränderung benötigt wird). Dort, wo man eine hohe Überdeckung hat, wählt man das Gitter enger; tief in der Erde, wo die Überdeckung geringer ist (d.h. das Gestein von weniger Strahlen durchlaufen wurde), wählt man ein entsprechend gröberes Gitter, d.h. lässt Zeilen und Spalten weg. Das hat den Vorteil, dass es erstens nicht mehr Bereiche als Messwerte gibt (im obigen Beispiel hieße das, die Bereiche wären weiter unterteilt, so dass durch manche Bereiche nur noch ein Strahl oder gar keiner mehr geht – mit der Folge, dass die Absorptionseigenschaften von manchen Bereichen nur ungenau bestimmt sind, d.h. die Eigenschaften der Bereiche untereinander eine starke Abhängigkeit haben), und zweitens bei der Berechnung weniger Stützstellen berücksichtigt werden müssen. Das theoretische Auflösungsvermögen der seismischen Tomographie wird von der Güte des Modells, d.h. von der Anzahl der Stützstellen (je mehr, desto besser) als auch von ihrer Abhängigkeit untereinander (je unabhängiger, desto besser), bestimmt. Das Einfügen zusätzlicher Stützstellen alleine bringt nichts, weil sich dadurch ihre Abhängigkeit verstärkt.

Wie lässt sich die Abhängigkeit der Stützstellen berechnen? Wenn man in obiger Gleichung:

$$\Delta\vec{f} = \left(s_{kl}^T s_{lk} + \theta^2 \mathbf{1} \right)^{-1} s_{kl}^T \Delta\vec{T}$$

einsetzt:

$$\Delta\vec{T} = s_{lk} \Delta\vec{f}_w$$

ergibt sich:

$$\Delta\vec{f} = \underbrace{\left(s_{kl}^T s_{lk} + \theta^2 \mathbf{1} \right)^{-1} s_{kl}^T s_{lk}}_R \Delta\vec{f}_w$$

R ist die *Auflösungsmatrix*, die angibt, wie stark die Werte einer Stützstelle von denen der Nachbar-Stützstellen abhängen; es hat sich bewährt, das Gitter so weit auszudünnen, dass alle Stützstellen einen R -Wert von mindestens 0,2 haben.

Eine Alternative zu irregulären Gittern sind Voronoi-Polygone.

Das maximale Auflösungsvermögen ist gleich einer Fresnel-Zone, d.h. gleich eines Bereichs, in dem der Wegunterschied eine halbe Wellenlänge ist:

$$r = \sqrt{\left(z + \frac{\lambda}{4} \right)^2 - z^2} = \sqrt{\frac{2\lambda z}{4} + \frac{\lambda^2}{16}} \approx \sqrt{\frac{1}{2}\lambda z} = \frac{1}{2}\sqrt{\lambda l}$$

7 Kontrolle des Modells: Varianz

Die *Datenvarianz* gibt an, wie groß der Unterschied zwischen den gemessenen Daten und den berechneten Modelldaten ist. Dazu steht im Zähler die Summe über die gewichteten (Koeffizient w) quadratischen Residuen, und im Nenner die Differenz aus der gewichteten Anzahl der Messungen und der Anzahl der invertierten Parameter:

$$VAR(data) = \frac{\sum_{n=1}^l w_n (t_n^{obs} - t_n^{cal})^2}{\left(\sum_{n=1}^l w_n\right) - m_{inv}}$$

Die *Lösungsvarianz* gibt die Modelländerung an, d.h. wie stark man das Modell mit Hilfe von Dämpfung usw. korrigiert hat:

$$VAR(solution) = \frac{\sum_{x=1}^{n_x} \sum_{y=1}^{n_y} \delta_i (v_{z,xy} - \bar{v}_z)^2}{m_{inv,z}}$$

Ziel ist es, einen Kompromiss zu finden, bei dem beide Varianzen gering gehalten werden können. Je stärker man mit Hilfe der Dämpfung die Datenvarianz minimiert, desto größer wird die Lösungsvarianz, d.h. desto weniger hat das ursprüngliche theoretische Modell mit dem verwendeten Modell zu tun – und umgekehrt.

8 Raytracing-Verfahren

Bei Inversionsverfahren (d.h. Rückwärtsproblemen) wie der seismischen Tomographie muss ein mögliches Modell berechnet werden – in diesem Fall der Weg des Strahls vom Schusspunkt bis zum Empfänger. Dies erledigen Raytracing-Verfahren, die iterativ die Laufzeit eines Strahls berechnen und anschließend den Laufweg anpassen, bis sich nach mehreren Durchgängen schließlich eine Laufzeit ergibt, die mit der gemessenen Laufzeit gut übereinstimmt. Folgende Prinzipien kommen zum Einsatz:

- **Shooting:** Hier wird vom Schusspunkt ausgehend der Laufweg eines Strahls berechnet, und die Startrichtung des Strahls so lange verändert, bis er am richtigen Endpunkt ankommt. Dadurch kann man auch ausrechnen, wie stark die geometrische Amplitudenabnahme ist, weil sich aus den durchgerechneten Laufwegen ergibt, wie stark sich der Bereich, in dem die Strahlen ankommen, mit der Variation des Abstrahlwinkels verändert, d.h. auf welchen Raumwinkelbereich die Energie verteilt wird. Nachteilig bei diesem Verfahren ist, dass es sehr viel Rechenleistung benötigt.
- **Bending:** Hier hält man Schusspunkt und Geophonpunkt fest und biegt den Strahl so lange zurecht, bis die Laufzeit passt. Dadurch bekommt man immer eine Lösung (weil der Endpunkt des Strahls per Voreinstellung gleich dem Geophonpunkt ist), die aber u.U. falsch sein kann (es wird ein lokales Laufzeitminimum statt des globalen gefunden), und die Amplitude kann nicht modelliert werden. Dafür benötigt das Verfahren weniger Rechenleistung.
- Graphische Methoden
- Finite Differenzen