Lecture 3: The convolutional model and standard seismic processing

Practical velocity estimation = global optimization by hand

Convolutional model = linearized, high frequency approximation to predicted seismogram:

$$DS[c]\delta c(h,t;x_m) \simeq w *_t (A(\bar{\zeta}(h,t),h)R(\bar{\zeta}(h,t)))$$

$$R(z) = \frac{1}{2}\frac{d}{dz}\frac{\delta c(z)}{c(z)}$$

$\bar{\zeta}$ is for each half-offset h the inverse function of the two-way traveltime $\bar{\tau}$:

$$\bar{\zeta}(2\bar{\tau}(z,h),h) = z$$

In the quarter-plane $r, z \geq 0$, $\bar{\tau}$ satisfies the 2D eikonal

$$\left(\frac{\partial \bar{\tau}}{\partial z}\right)^2 + \left(\frac{\partial \bar{\tau}}{\partial r}\right)^2 = \frac{1}{c(z)^2}$$

Most of standard seismic data processing is based on this very simple approximation! Standard reference: Yilmaz, 1987.

Given c it's easy to construct a least squares estimate of δc, or it would be if not for $w*_t$.

Fact: w is *bandlimited*, i.e. does not define an invertible convolution operator.

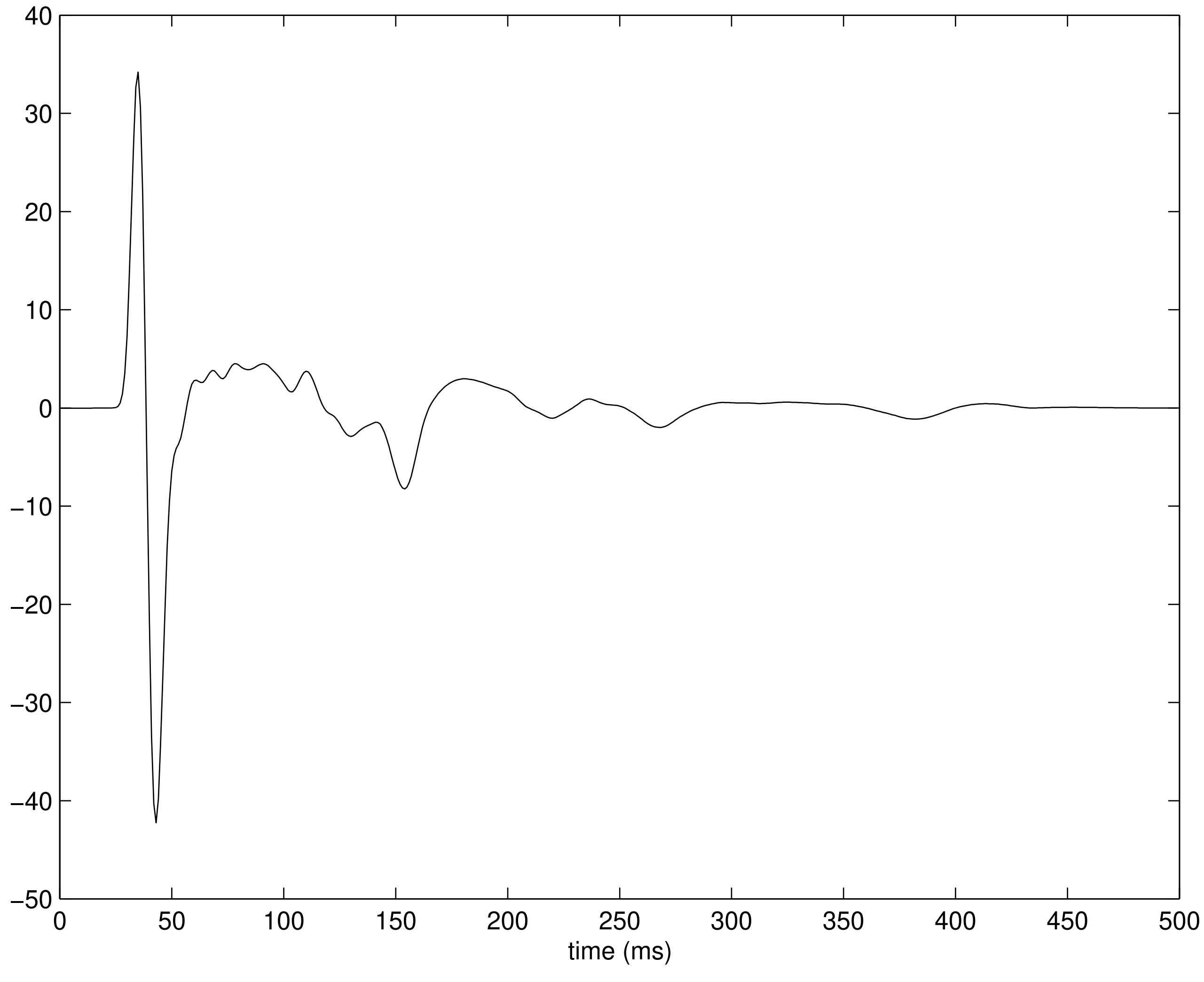
Fiction: $(w*_t)^{-1}$ exists in an effective enough sense that we can apply it to the data and thereafter treat $w(t) \sim \delta(t)$

Truly Awkward Fact: $w(t)$ is not known *a priori* - must be estimated from data, or measured (but measurement is surprisingly hard!)

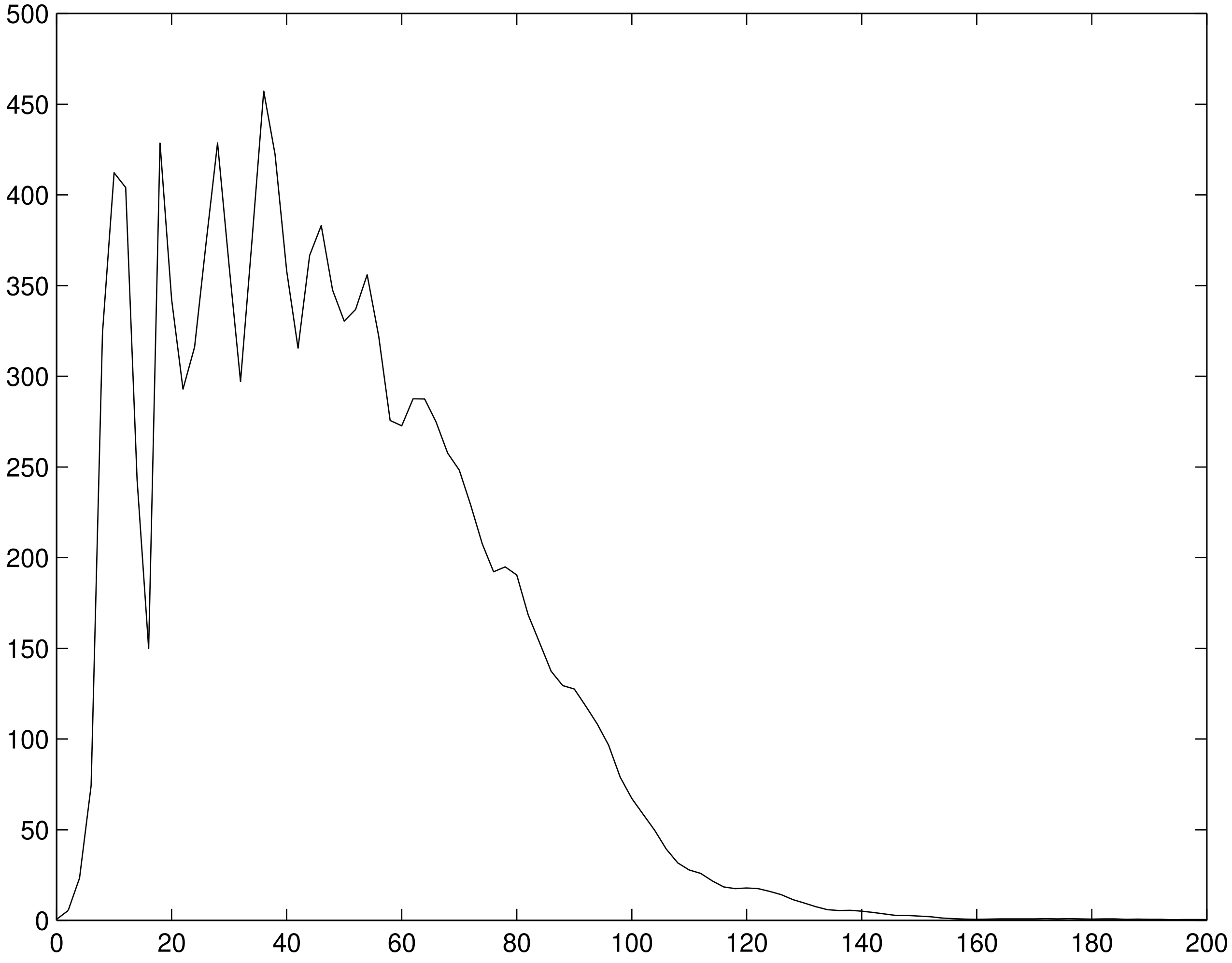
From now on treat $w \sim \delta$, even though it is at best a bandpass filter. Thus

$$DS[c]\delta c(h,t;x_m) \simeq (A(\bar{\zeta}(h,t),h)R(\bar{\zeta}(h,t)))$$

Reasonable research problem: understand how the fact that $w \neq \delta$ works its way through the steps to follow.



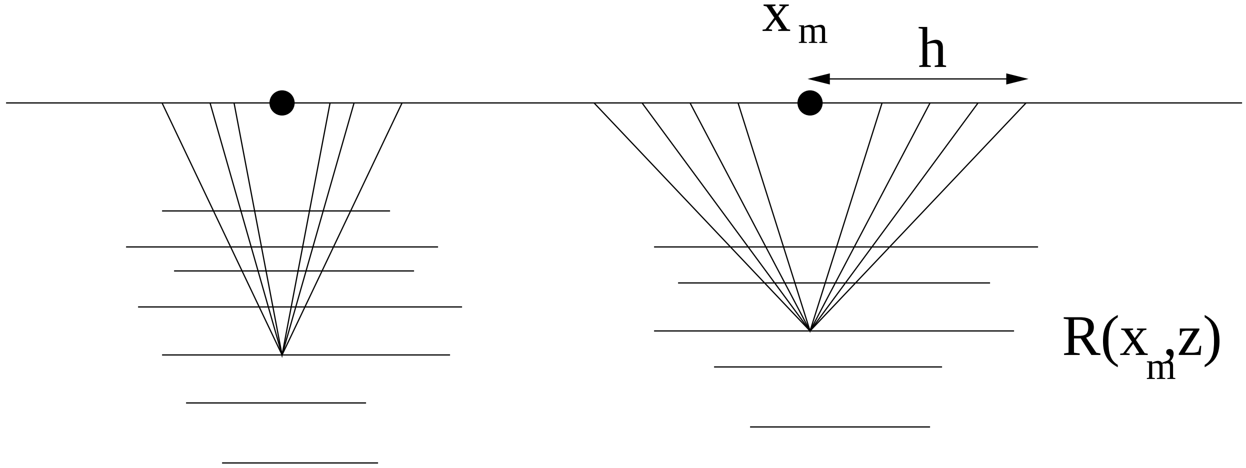
Source pulse for Mobil AVO survey, measured in deep water using hydrophone hanging 250 m below airgun array.



Power spectrum of Mobil AVO source pulse.

Now for the answer to yesterday's puzzle: why does x_m occur on the LHS but not on the RHS?

In fact it needs to be in both places! Layered approximation in *standard seismic processing*: Earth is layered, but different layers at different places. Each point on the Earth's surface is (potentially) a midpoint; seismic traces with this midpoint contribute to the estimation of velocity and reflectivity on the "virtual well" below this point, determines $c = c(x_m, z), R = R(x_m, z)$. Problems for various midpoints x_m independent $\Rightarrow$ concentrate on generic *midpoint gather*, suppress x_m from notation.



Least squares artifacts: the gradient is

$$\nabla J[c, \delta c; S^{\mathrm{obs}}] = DS[c]^*(S[c] - S^{\mathrm{obs}})$$

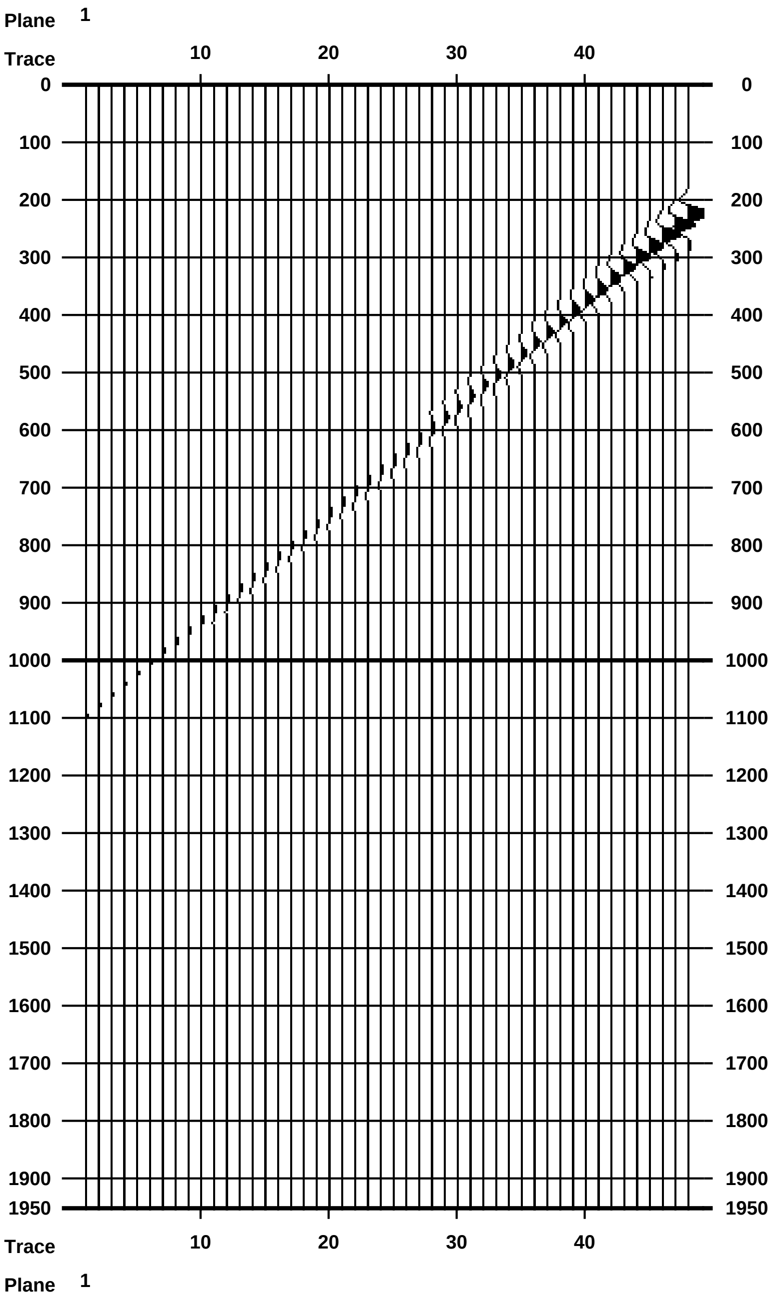
so we need to compute the (L^2) adjoint of $DS[c]$.

Simplify the computations in two steps:

(i) The relation $\delta c \to R$ is simple, so from now until further notice regard the seismogram as depending on R (so the model is now (c, R)):

$$DS[c]\delta c(h,t) = F[c]R(h,t) \equiv A(h, \bar{\zeta}(h,t))R(\bar{\zeta}(h,t))$$

(ii) From geometric optics, the "reference" seismogram $S[c]$ is all forward scattering! The signals caused by nonzero fluctuations δc or R typically occur *after* the energy in $S[c]$. Standard practice: construct a *mute* $\Phi(h,t)$, i.e. cutoff function, which $= 1$ where reflections are strong and $= 0$ on the support of $S[c]$ (see figure). More on construction of mute later.



Reference seismogram $S[c]$, $c = c(z)$ as in smoothness tests (lecture 2).

Mean square error of muted seismogram:

$$J[c,R;S^{\mathrm{obs}}] = \frac{1}{2}\|\Phi(S[c+\delta c] - S^{\mathrm{obs}})\|^2$$

$$\simeq \frac{1}{2}\|\Phi(F[c]R - S^{\mathrm{obs}})\|^2$$

$$\nabla_R J[c,R;S^{\mathrm{obs}}] = F[c]^*\Phi^2(F[c]R - S^{\mathrm{obs}})$$

Least squares estimate of R given c, S^{obs}:

$$R \sim (F[c]^*F[c])^{-1}F[c]^*S^{\mathrm{obs}}$$

Computation of adjoint:

$$< (\Phi F[c])^* e, R >=< e, \Phi F[c] R >$$

$$= \sum_h \int dt\, e(h,t)\Phi(h,t)A(h,\bar{\zeta}(h,t))R(\bar{\zeta}(h,t))$$

$$= \int dz \left(\sum_h (e\Phi)(h, 2\bar{\tau}(h,z))\left(2\frac{\partial\bar{\tau}}{\partial z}(h,z)\right)A(h,z)\right)R(z)$$

$$=< \Sigma \mathcal{F}[c]^* \Phi e, R >$$

where

- Σ is the *stacking operator*

$$\Sigma\phi(z) = \sum_h \phi(h,z)$$

  whose adjoint is the "regard function of z as function of h, z" or *spreading operator*;

- $\mathcal{F}[c]$ is the *incoherent* version of $F[c]$, whose domain is some appropriate space of *incoherent reflectivity* functions $\mathcal{R}(h,z)$ of h *and* z:

$$\mathcal{F}[c]\mathcal{R}(h,t) = A(h,\bar{\zeta}(h,t))\mathcal{R}(h,\bar{\zeta}(h,t))$$

  and whose adjoint is

$$\mathcal{F}[c]^* e(h,z) = 2A(h,z)\frac{\partial\bar{\tau}}{\partial z}(h,z)e(h,2\bar{\tau}(h,z))$$

Note that $F[c] = \mathcal{F}[c]\Sigma^*$, so

$$(\Phi F[c])^*(\Phi F[c])R(z) = \Sigma\mathcal{F}[c]^*\Phi^2\mathcal{F}[c]\Sigma^* R(z)$$

$$= \left(\sum_h 2\Phi^2(h, 2\bar{\tau}(z,h)) A(h,z)^2 \frac{\partial \bar{\tau}}{\partial z}(h,z) \right) R(z)$$

This mild-mannered formula has momentous consequences:

- The normal operator is multiplication by a smooth function.
- Inverse is easy to construct - multiply by reciprocal.
- Also any zones of rapid change or oscillations (or singularities) in R are preserved by normal operator.
- As shown in Lecture 2, adjoint is result of time-reversal process. If data is *consistent*, then application of adjoint *refocuses* reflectivity = produces large oscillations in same places as original, though perhaps with different amplitude. Will reinterpret this property mathematically, later.

- adjoint and (least squares) inverse differ only by a smooth scale factor, therefore have same singularities or zones of rapid change.
- (Geo)physical significance: if all you care about is *structure* = location of zones of rapid change, then application of adjoint reveals these - and you can even change the amplitude factor in the adjoint without affecting the structure of the output.!

Standard processing uses a simpler operator in place of least squares inverse or adjoint - outputs a field with same structure as least squares inverse:

$$M[c]e(z) = \sum_h e(h, 2\bar{\tau}(h,z))$$

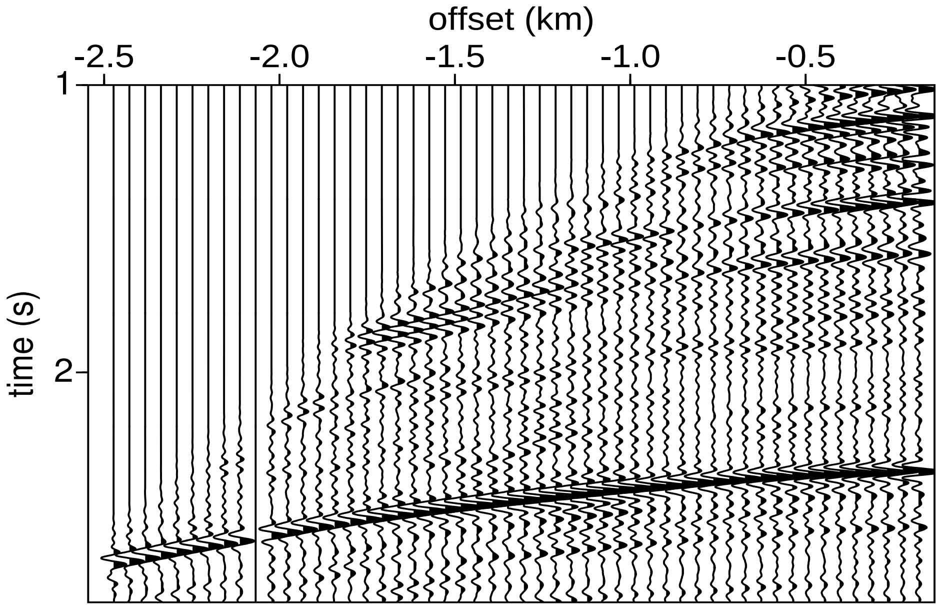
$$= \Sigma \mathcal{M}[c]e(z)$$

$$\mathcal{M}[c]e(h,z) = e(h, 2\bar{\tau}(h,z))$$

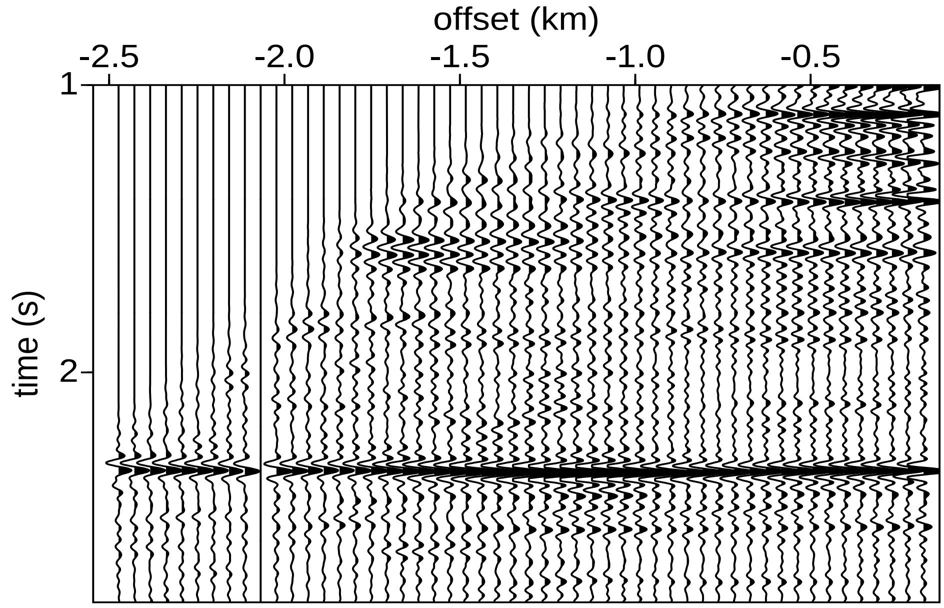
Normal Moveout Correction or *NMO* operator (depth version). "NMO, stack" (i.e. application of $M[c]$) for each midpoint. c may also depend on x_m. Upshot: "stacked trace" for each x_m. Aggregate = "stacked section", image of Earth reflectivity.

Expectation: where Earth conforms to layered model, structure should appear accurately in stacked section. [Significantly nonlayered structure has no reason to appear accurately, and doesn't! Getting it right is the job of migration - later.] Where constant density acoustic model accounts for most of data, rescaling will accurately recover $R(z)$.

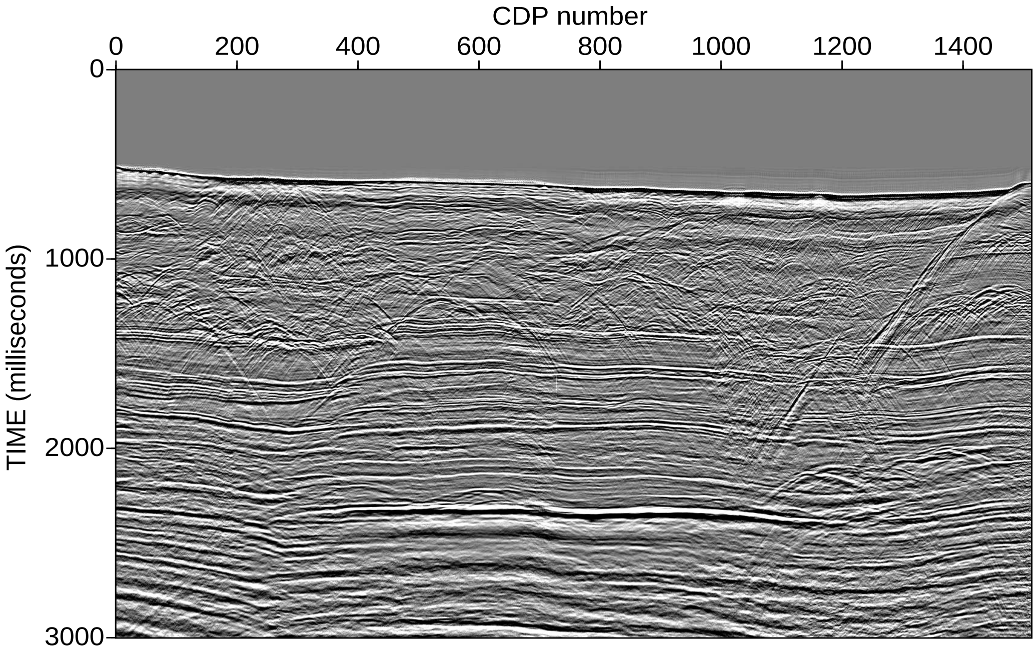
Example: marine survey from Mississippi Canyon region of Gulf of Mexico, courtesy Exxon Prod. Res. Co. Smooth background c extracted by process to be revealed later.



Typical midpoint gather.



After application of NMO operator



After stack, 1001 midpoints = stacked section

All of this depends on having the (?) right smooth velocity c. How do you find it?

Velocity Estimation Concept #1. Least squares: find R **and** c to minimize

$$J[c,R;S^{\mathrm{obs}}] \simeq \frac{1}{2}\|\Phi(F[c]R - S^{\mathrm{obs}})\|^2$$

Since we already know how to estimate R given c, remove it from contention:

$$R[c] \equiv ((\Phi F[c])^*(\Phi F[c]))^{-1}(\Phi F[c])^* S^{\mathrm{obs}}$$

Reduced least squares objective:

$$J^{\mathrm{red}}[c;S^{\mathrm{obs}}] = \frac{1}{2}\|\Phi(F[c]R[c] - S^{\mathrm{obs}})\|^2$$

then charge ahead and optimize J^{red} using eg. quasi-Newton method.

There are two things wrong with this program:

- takes acoustic model, approximations too literally - likely to leave large residuals
- it won't work - know this already from "layer over half space"

Velocity Estimation Concept #2.

Observe that F^*F computation $\Rightarrow$ *when velocity is correct, NMO corrected common midpoint ("CMP") gather is "flat": locations of zones of rapid change are independent of h.*

Suppose that data is consistent:

$$S^{\mathrm{obs}}(h,t) = A^*(h,\bar{\zeta}^*(h,t))R^*(\bar{\zeta}^*(h,t)),$$

where $\bar{\zeta}^*(h,t)$ pertains to $c^*(z)$. Then

$$\mathcal{M}[c]S^{\mathrm{obs}}(h,z) = A^*(h,\psi(h,z))R^*(\psi(h,z)),$$

$$\psi(h,z) = \bar{\zeta}^*(h,\bar{\tau}(h,z))$$

A^* smooooth $\Rightarrow$ R^* determines locii of rapid change $\Rightarrow$ these are $h-$independent ("flat") if $\psi(h,z) = z$. Clearly $\psi(h,z) = z$ if $c = c^*$, but this seems unlikely if c differs "substantially" from c^*.

Moveout = time/offset or depth/offset relation on a constant phase surface

Semblance Hypothesis: if velocity estimate is substantially wrong, the NMO corrected CMP gather will exhibit *residual moveout.*

Can either assess residual moveout visually, or employ a numerical measure - seismologists do both. Standard measures of residual moveout: *stack power*

$$P[c,S^{\mathrm{obs}}] = \|M[c]S^{\mathrm{obs}}\|^2$$

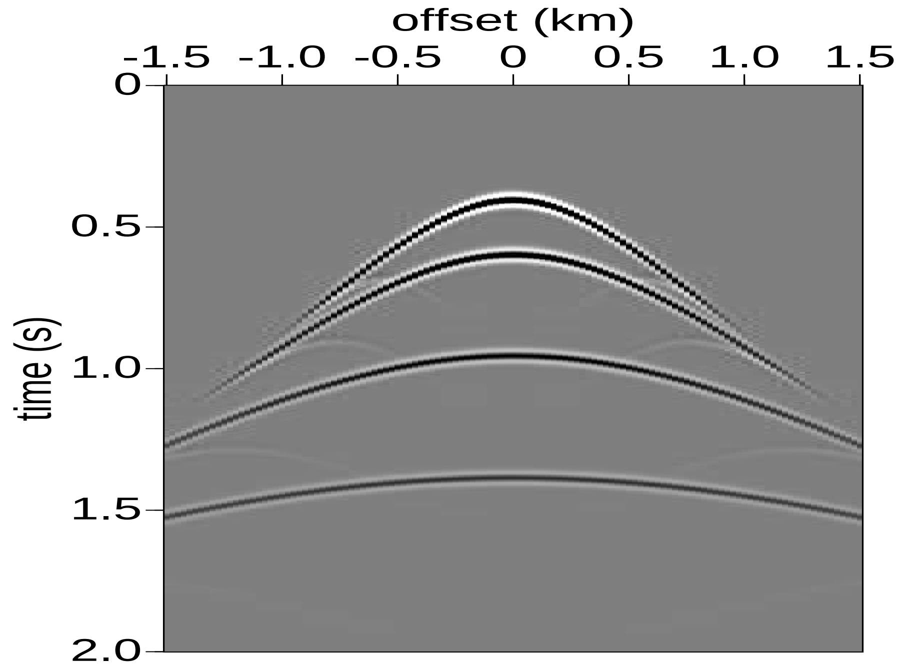
$$= \int dz\,(\sum_h \mathcal{M}[c]S^{\mathrm{obs}}(h,z))^2$$

and its nondimensionalized version, *semblance*

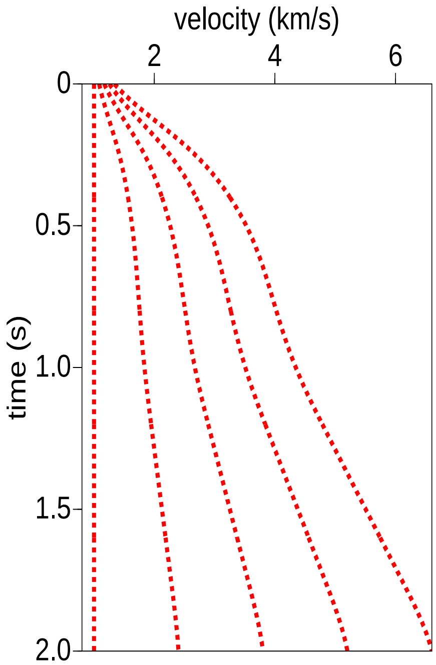
$$Q[c,S^{\mathrm{obs}}] = \frac{P[c,S^{\mathrm{obs}}]}{\|\mathcal{M}[c]S^{\mathrm{obs}}\|^2}$$

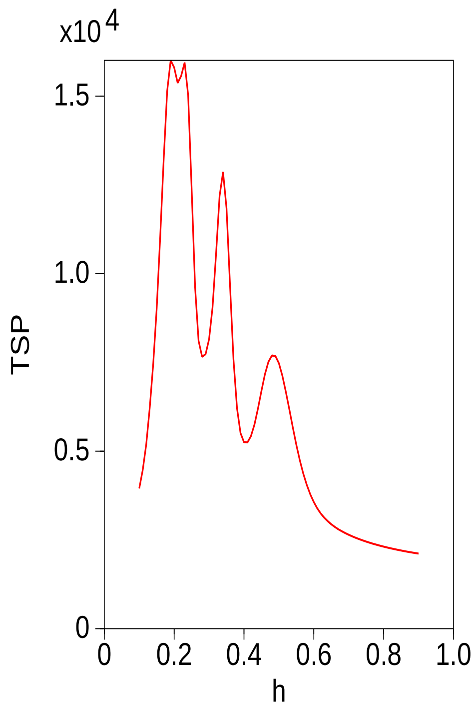
$$= \frac{\int dz\,(\sum_h \mathcal{M}[c]S^{\mathrm{obs}}(h,z))^2}{\int dz\,\sum_h(\mathcal{M}[c]S^{\mathrm{obs}}(h,z))^2}$$

Stack power and semblance exhibit "saturation" property similar to LS, so also not suited to descent-based optimization.



Example: synthetic "split spread" seismogram, c varies linearly from 1 to 5 km/s.





Left: samples along line segment in velocity space (*not* including target!)
Right: stack power along this line segment

Spurious maxima occur generically (Chauris, 1998) - global optimization a must.

Q: Is this an intractable optimization problem?

A: No, because it can be *diagonalized* i.e. decomposed into a product of (still global but) one dimensional optimization problems.

Diagonalization based on *hyperbolic moveout* approximation: for small offsets h,

$$\bar{\tau}(h,z) = \sqrt{\bar{\tau}(0,z)^2 + \frac{h^2}{\bar{c}(z)^2}} + O(h^4)$$

where

$$\bar{c}(z) = \sqrt{\frac{\int_0^z \, dz' \, c(z')}{\int_0^z \, dz' \, \frac{1}{c(z')}}}$$

Proof: Differentiating the eikonal equation for $\bar{\tau}$ twice with respect to h, get

$$\frac{\partial \bar{\tau}}{\partial z}\frac{\partial^3 \bar{\tau}}{\partial z \partial^2 h} + \left(\frac{\partial^2 \bar{\tau}}{\partial z \partial h}\right)^2$$

$$+\frac{\partial \bar{\tau}}{\partial h}\frac{\partial^3 \bar{\tau}}{\partial^3 h} + \left(\frac{\partial^2 \bar{\tau}}{\partial h^2}\right)^2 = 0$$

The vertical line $h = 0$ is a ray, and the travel time along it is

$$\bar{\tau}(0, z) = \int_0^z \frac{dz'}{c(z')} \text{ so } \frac{\partial \bar{\tau}}{\partial z} = \frac{1}{c}$$

Note that $\bar{\tau}$ is *even* in h, so all odd order derivatives vanish at $h = 0$:

$$\frac{1}{c}\frac{\partial^3 \bar{\tau}}{\partial z \partial^2 h}(0, z) + \left(\frac{\partial^2 \bar{\tau}}{\partial h^2}(0, z)\right)^2 = 0$$

This ODE for $\frac{\partial^2 \bar{\tau}}{\partial h^2}$ has the solution

$$\frac{\partial^2 \bar{\tau}}{\partial h^2}(0, z) = \frac{1}{\text{const.} + \int_0^z dz'\, c(z')}$$

Since $\bar{\tau}$ is asymptotic to $r/c(0)$, the constant must vanish. Upshot:

$$\bar{\tau}(h, z) = \bar{\tau}(0, z) + \frac{h^2}{2 \int_0^z dz'\, c(z')} + O(h^4)$$

whence

$$\bar{\tau}(h,z)^2 = \bar{\tau}(0,z)^2 + \frac{h^2\bar{\tau}(0,z)}{\int_0^z dz'\,c(z')} + O(h^4)$$

Take square roots and insert the relation between $\bar{\tau}(0,z)$ and $c(z)$. q.e.d.

Exercise: This formula is not in a form that geophysicists would recognize. To recover that, change variables $z \mapsto \bar{\tau}(0,z)$, with inverse change of variables $t \mapsto \bar{\zeta}(0,t)$. Write

$$t_0 = \tau(0,z),\ V_{\mathrm{int}}(t_0) = c(\bar{\zeta}(0,t_0)),$$

$$V_{\mathrm{RMS}}(t_0) = \bar{c}(\bar{\zeta}(0,t_0))$$

Then moveout curve through t_0 at $h=0$ is

$$t(h,t_0) \equiv \tau(h,\zeta(h,t_0)) = \sqrt{t_0^2 + \frac{h^2}{V_{\mathrm{RMS}}(t_0)^2}},$$

$$V_{\mathrm{RMS}}(t_0)^2 = \frac{1}{t_0}\int_0^{t_0} dt V_{\mathrm{int}}(t)^2$$

hence "RMS" notation; $v > 0 \Rightarrow V_{\mathrm{RMS}}$ tends to increase

What's been gained: the moveout curve through t_0 at $h=0$ depends on only one number $V_{\text{RMS}}(t_0)$, rather than on $c(z), 0 \le z \le \bar{\zeta}(0,t_0)$. Therefore the *local stack power*

$$\mathcal{P}(t_0) = (\sum_h S^{\text{obs}}(h, t(h,t_0)))^2$$

depends only on $V_{\text{RMS}}(t_0)$, *so you can optimize it by global search in 1D!*. Moreover, apart from a Jacobian factor, the global stack power function is the sum of the local ones:

$$P[V_{\text{RMS}}, S^{\text{obs}}] = \int dt_0 \, \mathcal{P}(V_{\text{RMS}}(t_0), t_0)$$

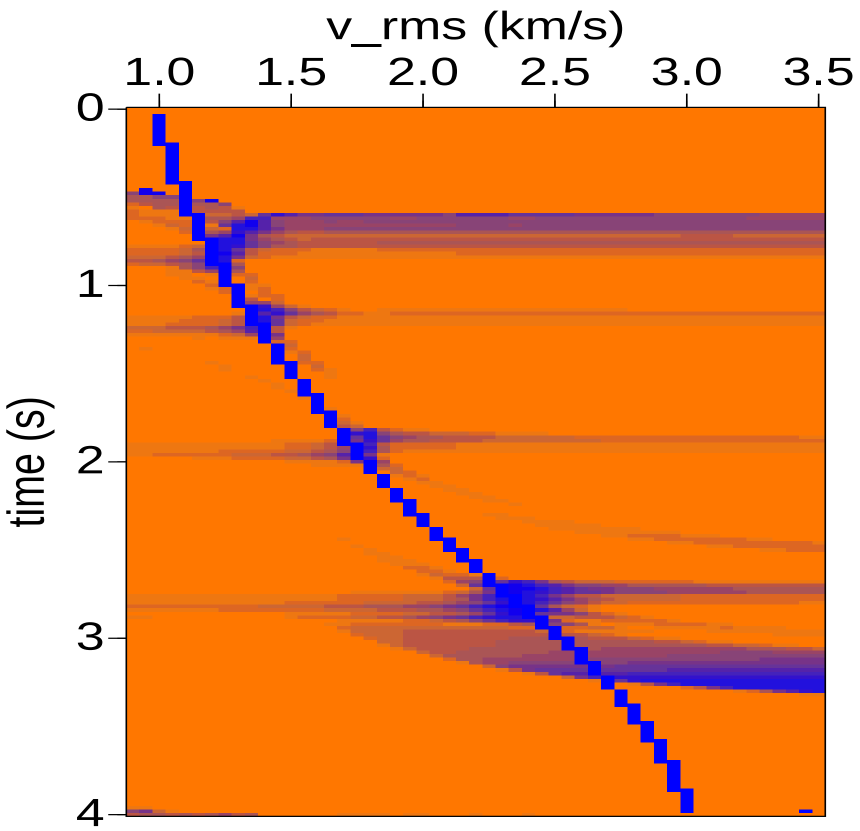
So you can globally optimize P (or Q) by carrying out the suite of 1D global optimizations of $\mathcal{P}$ (or $\mathcal{Q}$) over each $V_{\text{RMS}}(t_0)$, then extract $c(z)$.

Practicality: to make the behaviour of $\mathcal{P}$ a bit more stable, *sample* P, V_{RMS} at rate Δt_0 and sum stack power over a window:

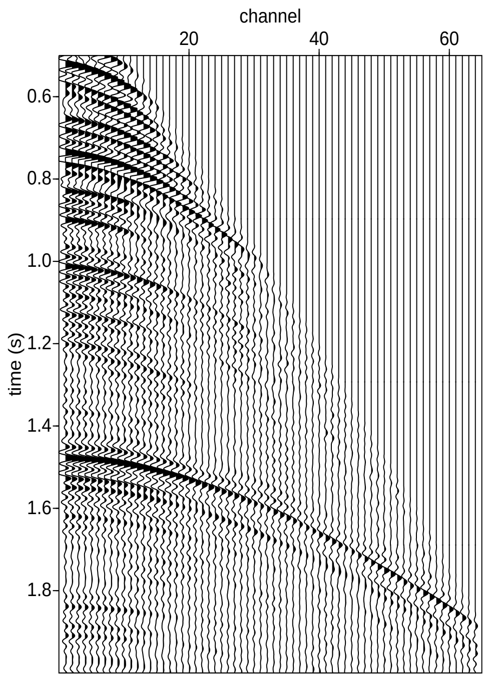
$$\mathcal{P}(V_{\text{RMS}}(t_0), t_0) = \int_{t_0-\Delta t_0}^{t_0+\Delta t_0} dt_0' (\sum_h S^{\text{obs}}(h, t(h,t_0')))^2$$

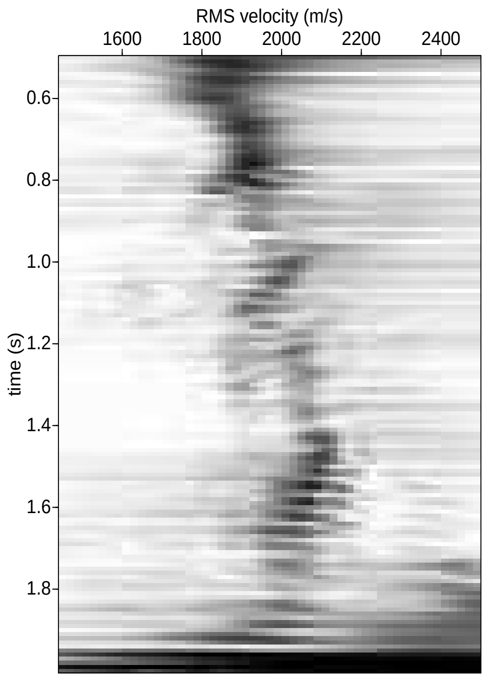
Standard display on which to carry out global optimization (visually!): *velocity spectrum* (Taner & Kohler, 1969) - contour or grey/color scale plot of $\mathcal{P}(V_{\mathrm{RMS}}(t_0), t_0)$ versus V_{RMS}, t_0.

Example: synthetic split spread



Example: Shell North Sea data



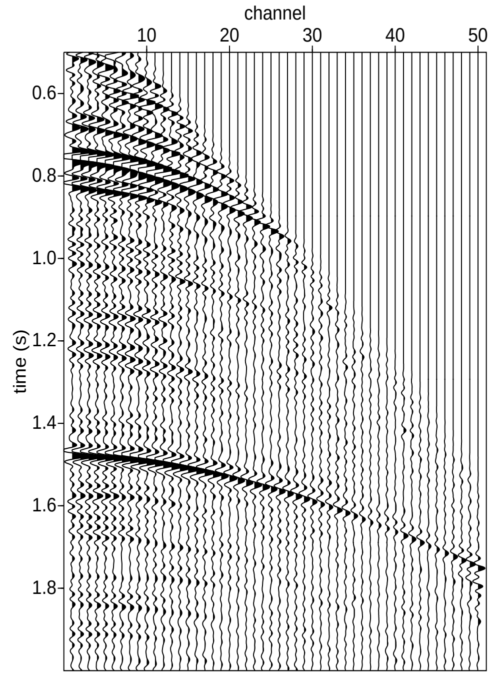


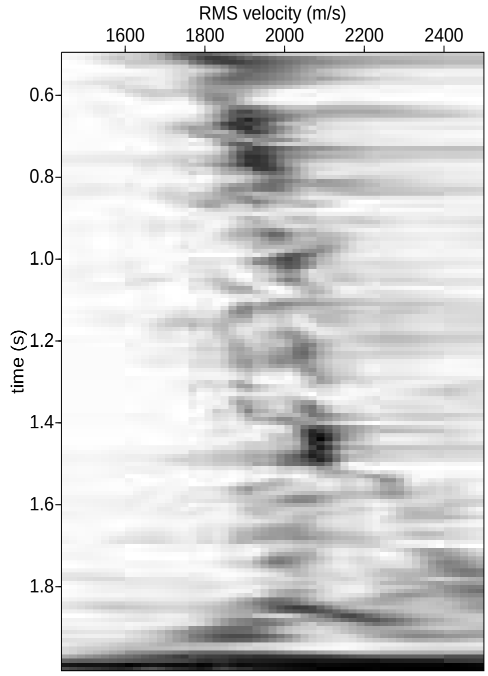
Left: a CMP gather

Right: velocity spectrum - general trend clear but details inconsistent (expect gentle increase) - which peaks to pick?

Where (some of) the multiple maxima arise: *multiple reflections* - convolutional model does not predict them.

Example: Shell North Sea data after Predictive Deconvolution = signal processing technique to suppress multiple reflections.





Left: CMP gather
Right: velocity spectrum - multiple peaks largely removed

Multiple source: sea bottom - 50 m @ 1500 m/s $\times$ 2 $\sim$ 0.07 s = lag time of first multiple reflection at small h.

Upshot: velocity spectrum and similar techniques based on hyperbolic moveout approximation and stack power or semblance work well when layered medium and linearization assumptions are not violently wrong:

- Earth structure is genuinely close to layered

- multiple reflections are not terrifically energetic, or removeable by signal processing

These conditions are met in some locations (parts of Gulf of Mexico, North Sea,...) but not in others (parts of G of M, North Sea,...!).

No effective analog of hyperbolic moveout approximation exists for traveltime in nonlayered material $\Rightarrow$ estimation of nonlayered velocities cannot be broken down into 1D global optimizations! $\Rightarrow$ least squares or stack power or...is a high dimensional global opt problem $\Rightarrow$ computationally intractable, must seek other approaches.