DATA RESTORATION WITH PARSIMONIOUS MODELS

Jon F. Claerbout

To be parsimonious, a model must have an abundance of zero or small values. Whether a given mathematical function of geophysical data represents a parsimonious model is a matter that depends largely on geological circumstances. If a parsimonious model exists, and the mathematical function is invertible, then the function can be utilized to extrapolate and interpolate an incomplete dataset.

Those of us who have worked on migrating stacked sections have been spoiled. Most geophysical datasets, such as *CDP* gathers, are seriously truncated, aliased, and gapped.

Examples of Parsimonious Models

1. $P(t=0,y,z)$. The migrated seafloor should have no echos above it. The suppression of small values above the seafloor in migrated data should be a good criterion for the horizontal interpolation of data near the seafloor. Likewise, off the side edge of a migrated section, depressing the semicircles leads to extrapolated values in the unmigrated section.
2. $P(\omega,k_y)$. An entire zone of (ω,k_y)-space should vanish by evanescence. Good quality data often has a narrowly banded dip spectrum so that most energy lies along lines in (ω,k_y)-space. For example, you should be able to extrapolate a strong, consistent dip at the truncation (in y) of a zero-offset section by the criterion that the dip spectrum should be sharpened.
3. $P(h,t=0,z)$. Downward-continued *CDP* gathers should have energy clustering near zero offset, $h=0$, with small values elsewhere.

4. $P(h, t'=\tau, p_0)$. Downward-continued, slanted, *CMP* gathers with pre-critical reflections (Gonzalez and Claerbout, SEP-16, p. 181) should have a clustering at some $dh/d\tau$, which determines interval velocity.

5. $P(t=0, z, p)$. Downward-continued, *CMP* gathers with refractions (Clayton and McMechan, SEP-24, p. 33) should have a clustering at $v(z) = 1/p(z)$.

All the above mappings are linear or quasi-linear. There is no "picking" of the data. All are easily invertible.

Window Functions and Stretch Functions

A dogmatic way to treat prior information of supposed vanishing portions of model space is by windowing, that is, transforming data to model space, then zeroing the offending values. Returning to data space, interpolated and extrapolated data values will be found. The trouble is that where it is already known, the resulting data will be inconsistent, to some degree, with the original data. Restoring the original data, enables repeating the process. Iteration may never converge since the raw data may be inconsistent with the ideology of the prior information. This suggests using a succession of weakening tapers. For example, off the side of a migrated section it seems reasonable to use a linear taper weight, which drops to zero at the radius of the migration semicircle. Likewise, tapers could be used around the velocity-estimation models.

Sometimes the parsimony of models results more wholly from geological circumstances. Prior information about wave propagation may play no role. It may be that the earth model is dominated by just a few dips. We must depend, it seems, on the possible parsimony of the earth model if we are to interpolate sections that are spatially aliased. The practical approach here is to use a data-dependent window, in other words, a *stretch* function. Increase large values and decrease small ones. Then iterate.

The way to choose an appropriate mix between window functions and stretch functions must, at the present time, be regarded as practitioner's art.

Definitions

r = raw data

x	=	missing data or "padding for r" or "tag" data (truncated, aliased, gapped)
M	=	parsimonious model
n	=	noise in the space of r caused by M being inconsistent with r
(r,x)	=	full dataset = raw data with padding
$L(r,x)$	=	function to generate a model from a full dataset
$L^{-1}(M)$	=	function to generate data $(r+n,x)$ from any model M
$\overline{M}$	=	$L(r,$ any $x)$ = a noise-free model
S	=	your choice of window and stretch functions

The minimum entropy stretch is defined by

$$S(M) = M\left[\frac{M^*M}{\ll M^*M \gg}\right]^{\varepsilon} \quad where \quad 1 \gg \varepsilon > 0$$

A Faster Algorithm

The model M proposed in each loop of this algorithm should be better than the model in the algorithm proposed in a previous paper (Extrapolation, Interpolation and Smoothing of Wave Fields), because more effort is made to get M consistent with r. As a result, you can start with bigger values of ε and hope for faster convergence.

Algorithm ("=" means "←")			Comments ("=" means "define")		
$\overline{M}$	=	$L(r,0)$	$\overline{M}$	=	noise-free model, zero pad
M	=	$S(\overline{M})$	M	=	noisy model, less entropy
Begin iteration loop.			Good luck!		
$(r+n,x)$	=	$L^{-1}(M)$	Discover noise n and revised pad x		
$\overline{M}$	=	$L(r,x)$	$\overline{M}$	=	noise-free, pad is same x
N	=	$M-\overline{M}$	N	=	Noise. Note M and $\overline{M}$ have same x.
M	=	$S(\overline{M})$	M	=	a new noisy model
P	=	$M-\overline{M}$	P	=	proposed perturbation to $\overline{M}$
P	=	$P-N\frac{(N \cdot P)}{(P \cdot P)}$	Remove noise, you hope.		
M	=	$\overline{M}+P$	M	=	proposed model
End iteration loop.					

Comments on Algorithm

Strength. With luck and experience, a big ε may do the job on one pass without iteration.

Weakness. Noise N may oscillate from one iteration to the next, so it may need to be averaged. No good ideas about appropriate size of ε. We rely on experience and sense.