Stanford Exploration Project, Report 103, April 27, 2000, pages 215–330

Short Note

Random lines in a plane

Jon Claerbout[1]

INTRODUCTION

Locally, seismic data is a superposition of plane waves. The statistical properties of such superpositions are relevant to geophysical estimation and they are not entirely obvious.

Clearly, a planar wave can be constructed from a planar distribution of point sources. Contrariwise, a point source can be constructed from a superposition of plane waves going in all directions. We can represent a random wave source either as a superposition of points or as a superposition of plane waves. Here is the question:

> Given a superposition of infinitely many impulsive plane waves of random amplitudes and orientations, what is their spectrum?

If you said the spectrum is white, you guessed wrong. Figure 1 shows that it does not even look white. It is lower frequency than white for good reason. Mathematically independent variables are not necessarily statistically independent variables.

RESOLUTION OF THE PARADOX

If we throw impulse functions randomly onto a plane, the power spectrum of the plane is the power spectrum of impulse functions, namely white.

Think of a 2-D Gaussian whose contour of half-amplitude describes an ellipse of great eccentricity. In the limit of large eccentricity, this Gaussian could be one of the lines that we sprinkle on the plane with random amplitudes and orientations. The spatial spectrum of such an eccentric Gaussian must be lower than that of a symmetrical point Gaussian because the spectrum along the long axis of the ellipsoid is concentrated at very low frequency.

Consider a single delta function along a line with an arbitrary slope and location in a plane. The autocorrelation of this dipping line is another dipping line with the same slope, but passing through the origin at zero lag. The polarity of the impulse function is lost in the autocorrelation; in the autocorrelation space, the amplitude of the dipping line is positive.

[1]**email:** claerbout@stanford.edu

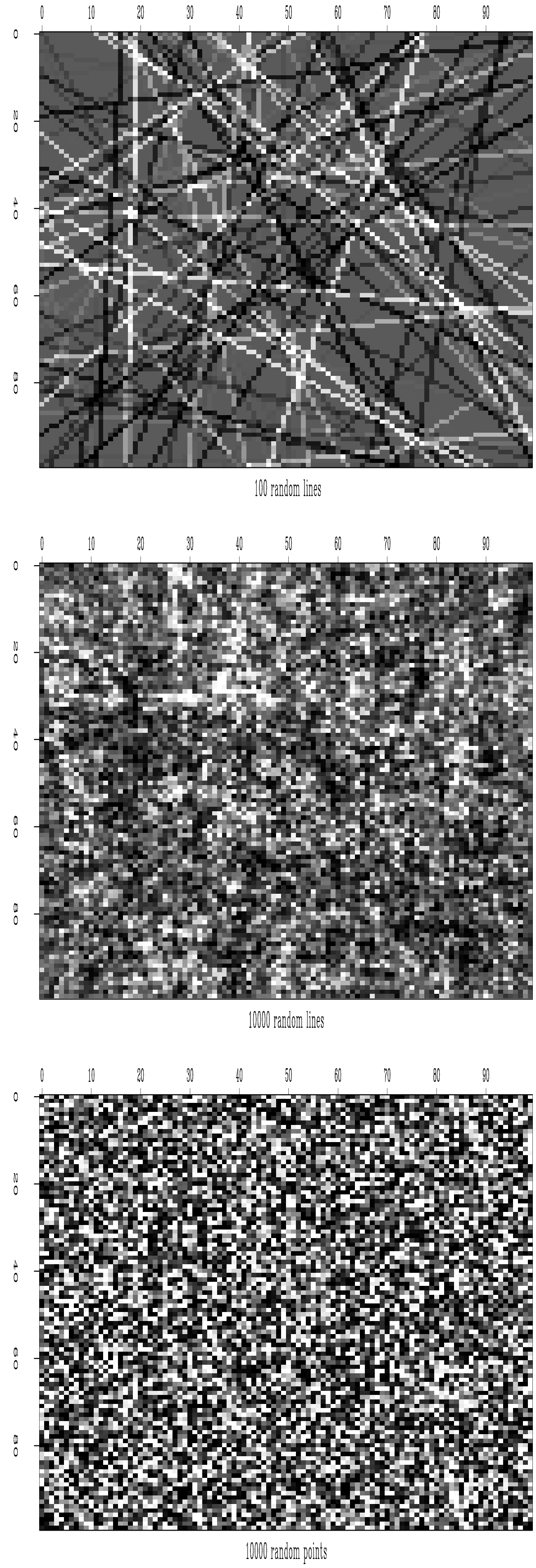

Figure 1: Top shows a superposition of 100 randomly positioned lines. Middle shows a superposition of 10,000 such lines. Bottom shows a superposition of 10,000 random point values. The bottom panel shows the most high frequency. Its spectrum is theoretically white. This paper claims that the middle panel is more representative of natural noises. jon3-randline [ER,M]

Now consider a superposition of many dipping lines on the plane. Its autocorrelation is the sum of the autocorrelations of individual lines. The autocorrelation of any individual line is a line of the same slope that is translated to pass through the origin. [The 2-D autocorrelation is not shown in the graphics here. You'll need to understand it from the words here. Sorry.] The autocorrelation is a superposition of lines of various slopes all passing through the origin, all having positive amplitude. This function would resemble a positive impulse function at the origin (and hence suggest a white spectrum). The function is actually not an impulse function, but, as we'll see, it is the pole $1/r$.

Consider an integral on a circular path around the origin. The circle crosses each line exactly twice. Thus the integral on this circular path is independent of the radius of the circle. Hence the average amplitude on the circumference is inverse with the circumference to keep the integral constant. Thus the autocorrelation function is the pole $1/r$.

FOURIER TRANSFORM OF 1/R

We would like to know the 2-D Fourier transform of $1/r$. Everywhere I found tables of 1-D Fourier transforms but only one place did I find a table that included this 2-D Fourier transform. It was at http://www.ph.tn.tudelft.nl/Courses/FIP/noframes/fip-Statisti.html

Sergey Fomel showed me how to work it out: Express the FT in radial coordinates:

$$\mathrm{FT}\left(\frac{1}{r}\right) = \int\int \exp[ik_x r\cos\theta + ik_y r\sin\theta]\,\frac{1}{r}\,r\;dr\;d\theta \tag{1}$$

$$\mathrm{FT}\left(\frac{1}{r}\right) = \int \delta[k_x\cos\theta + k_y\sin\theta]\;d\theta \tag{2}$$

To evaluate the integral, we use the fact that $\int\delta(f(x))dx = 1/|f'(x_0)|$ where x_0 is defined by $f(x)=0$ and the definition $\theta_0 = \arctan(-k_x/k_y)$.

$$\mathrm{FT}\left(\frac{1}{r}\right) = \frac{1}{|-k_x\sin\theta_0 + k_y\cos\theta_0|} \tag{3}$$

$$\mathrm{FT}\left(\frac{1}{r}\right) = \frac{1}{\sqrt{k_x^2+k_y^2}} = \frac{1}{k_r} \tag{4}$$

UTILITY OF THIS RESULT

At present we are accustomed to estimating statistical properties of seismic data by computing a 2-D prediction-error filter. This filter is needed to interpolate and extrapolate missing values.

Knowing that the prior spectral estimate is not a constant but instead is $1/k_r$ suggests a procedure that is more efficient statistically: By more efficient, I mean that a simpler model should fit the data, a model with fewer adjustable parameters.

INTERPOLATION

We'll need to know a wavelet in the time and space domain whose amplitude spectrum is $\sqrt{k_r}$ (so its power spectrum is k_r). Do not mistake this for the the helix derivative (Claerbout, 1998) whose power spectrum is k_r^2. What we need to use here is the square root of the helix derivative. Let the (unknown) wavelet with amplitude spectrum $\sqrt{k_r}$ be known as G.

1. Apply G to the data. The prior spectrum of the modified data is now white.
2. Estimate the PEF of the modified data.
3. The interpolation filter for the original data is now G times the PEF of the modified data.

Why is this more efficient? The important point is that the PEF should estimate the minimal practical number of freely adjustable parameters. If G is a function that is lengthy in time or space, then the PEF does not need to be.

How important is this extra statistical efficiency? I don't know.

WHAT ARE THE NEXT STEPS?

- Compute $\sqrt{k_r}$ in physical space and look at it. How best to do this? How best to package the software?
- Invent a synthetic data test.
- Think about how it might impact Sean (or Sergey). Which example of Sean's would be worth redoing?
- Extend this idea to 3-D (lettuce versus noodles).
- Matt Schwab and I were frequently disappointed in the performance of local PEFs for the task of visualizing data. This might explain it. Which example of his might be worth redoing?

REFERENCES

Claerbout, 1998, Multidimensional recursive filters via a helix: Geophysics, **63**, 1532–1541.