Basics of blocky function estimation

Consider vertical seismic propagation where we can ignore the source waveform, multiple reflections, shear waves, and the low frequence change of impedance and velocity with depth. We can include any or all these real world complications later, but for now we investigate blocky function estimation in its simplest context.

Let us use the admittance instead of the impedance (because we can use the letter y instead of i for the math). We will be estimating the admittance $y(z)$ as a function of the observed reflectivity $c(z)$ where z is the depth.

The basic physics defines reflectivity c as the relative change in admittance.

$$\frac{1}{y}\frac{dy}{dz} = c$$

which resembles the more familiar expressions for impedance and admittance like

$$c(z+1/2) = \frac{y_{z+1} - y_z}{y_{z+1} + y_z}$$

We linearize by allowing $y(z)$ to fluctuate only about a mean of unity $y(z) \approx 1$. The data fitting regression is

$$0 \quad \approx \quad r_d \; = \; \frac{dy}{dz} - c(z)$$

Since blocky functions have zero slope (except for jumps between blocks) we use the regularization regression equation

$$0 \quad \approx \quad r_m \; = \; \frac{dy}{dz}$$

We change variables from the blocky variable $y(z)$ to a sparse variable $s(z)$ by the definition $dy/dz = s$. This is the usual preconditioning the regularization. The resulting pair is

$$0 \;\approx\; r_d(z) \;=\; s(z) - c(z) \qquad (1)$$

$$0 \;\approx\; r_m(z) \;=\; s(z) \qquad (2)$$

Notice each value of z is an independent problem. How do we arrange for $s(z)$ to come blocky from the regression? The answer is to use a sharp bottomed penalty function like $|r|$ for the regularization and a flater bottomed penalty like r^2 for the data fitting. Why does this work? The absolute value function can dominate the square function only near the origin, only when $|c|$ is small.

If you seek rough comprehension of why this works, instead of using $L2$ for the data try the flat-bottomed penalty function $N(r) = \max(0, |r| - r_T)$ where r_T is a threshold on the data fitting. There is an epsilon such that for small data values $|c| << r_T$ the solution s is completely captured by the regularization so $s = 0$. Otherwise (for positive data $c > r_T$) the solution tracks the data but a little below it $s = c - r_T$.

Why are we on the right track? It would be easy enough to come up with methods of blocky function estimation not based on minimizing convex functions. The advantage of our approach is that we get an answer that is unique, independent of the starting location, it tries to match both our goals, and being "of inverse theory" we know how to integrate it with more complicated models (such as shot waveform estimation).

PROPOSED CODE SKETCH

Define r_T as the transition threshold between L1 and L2 in the hybrid norm. To get blocky functions we could fit their derivative with L1 norm as the regularization and L2 for data fitting. Alternately we can use the hybrid norm for both regularization and data fitting but use a different threshold r_T for each. For example we might define the thresholds $r_B = 60\%$ for the data fitting $r_C = 95\%$ for regularization seeking spiky models (with blocky derivatives). Changing equations (1) and (2) to more familiar nomenclature we restate the earlier goals:

$$0 \approx r_d(z) = m(z) - d(z) \tag{3}$$

$$0 \approx r_m(z) = m(z) \tag{4}$$

$\mathbf{m} = \mathbf{d}/2$
loop over choosing r_B and r_C { # Priors? sorting over $d(z)$? dare sort over $m(z)$?
loop over time points {
loop over non-linear iterations {
Get 1st and 2nd derivatives of hybrid norm $B'(m-d)$ and $B''(m-d)$ for data goal.
Get 1st and 2nd derivatives of hybrid norm $C'(m)$ and $C''(m)$ for model goal.
Plan to find α to update $m = m + \alpha$
Taylor series for data penalty $E(r_d) = B + B'\alpha + B''\alpha^2/2$
Taylor series for model penalty $E(r_m) = C + C'\alpha + C''\alpha^2/2$

$$0 = \frac{\partial}{\partial \alpha}(E(r_d) + E(r_m))$$

Quiz: Compute α

$m = m + \alpha$
} end of loop over non-linear iterations
} end of loop over all time points
} end of loop over r_B and r_C estimations

I'd revise the definition of the hybrid norm to $\sqrt{r_T^2 + r^2}$. Then the gradient is asymtotically independent of r_T^2. I'd also start with epsilon$= \epsilon = 1/3$.