October 28, 2011

The magic of Logarithm method

Qiang Fu

ABSTRACT

Jon proposed a linear Fourier domain Bidirectional deconvolution method yesterday. Soon we find it does not work well. However, we learned a important lesson from the attempt to implement the linear method – what makes logarithm method magic.

For the logarithm Fourier domain method, the residual of the deconvolution is

$$r = \mathrm{FT}^{-1}\, D(Z(w)) e^{U(Z(w))} \tag{1}$$

$$r = \mathrm{FT}^{-1}\, D(Z(w)) e^{U(Z(w))}, \tag{2}$$

And the linear method residual is

$$r = \mathrm{FT}^{-1}\, D(Z(w)) F(Z(w)), \tag{3}$$

Let's look at the gradient of these two residuals,

Logarithm method gradient of residual:

$$\begin{gathered} \mathbf{r} = \mathrm{FT}^{-1}\, D(Z) e^{U(Z)} \\ \mathbf{r} = \mathrm{FT}^{-1}\, D(Z) e^{\cdots + u_2 Z^2 + u_3 Z^3 + u_4 Z^4 \cdots} \\ \frac{\partial \mathbf{r}}{\partial u_\tau} = \mathrm{FT}^{-1}\, D(Z) e^{\cdots + u_2 Z^2 + u_3 Z^3 + u_4 Z^4 \cdots} Z^\tau \\ \frac{\partial r_t}{\partial u_\tau} = r_{t+\tau} \end{gathered} \tag{4}$$

Linear method gradient of residual:

$$\begin{gathered} r = \mathrm{FT}^{-1}\, D(Z) F(Z) \\ r = \mathrm{FT}^{-1}\, D(Z)(\cdots + f_2 Z^2 + f_3 Z^3 + f_4 Z^4 \cdots) \\ \frac{\partial r}{\partial f_\tau} = \mathrm{FT}^{-1}\, D(Z) Z^\tau \\ \frac{\partial r_t}{\partial u_\tau} = d_{t+\tau} \end{gathered} \tag{5}$$

We find the gradient of linear method is a shifted data, but the gradient of logarithm method is a shifted residual. This is because in logarithm method, we have a term e^U, and $(e^x)' = e^x$, so this term will bring the decon filter into the gradient. That is why we get residual rather than data itself (data combined decon filter equals residual). This is the most important difference between these two methods.

The normal equation $N()$ for the deconvolution is the summation of the norm or penalty function $H()$ on every residual component. (I will refer norm or any non-norm penalty function as "norm" in later of this paper, without to distinguish the difference between norm and non-norm penalty function)

$$\mathrm{N}(\mathbf{r}) = \sum_t \mathrm{H}(r_t) \tag{6}$$

The gradient of the whole normal equation is gradient of residual multiply by the gradient of the norm with respect to the residual,

$$\begin{aligned} \Delta\mathbf{u} &= \frac{\partial\,\mathrm{H}(\mathbf{r})}{\partial\mathbf{u}} \\ \Delta u_\tau &= \frac{\partial \sum \mathrm{H}(r)}{\partial u_\tau} = \sum_t \frac{\partial r_t}{\partial u_\tau}\frac{dH(r_t)}{dr_t} \\ &= \sum_t \frac{\partial r_t}{\partial u_\tau} H'(r_t) \end{aligned} \tag{7}$$

If we substitute the gradient of these two residuals in to the gradient of the whole normal equation respectively, we can find gradient of the whole normal equation the gradient of the normal equation (update direction for residual) is data correlate with the first derivative of the norm . And for logarithm method it is residual correlate with the first derivative of the norm.

If we put $H() =$"L2 norm" here, the gradient of the normal equation will be very familiar for us.

We know when the iteration converged, the gradient of the normal equation should be vanished. That tell us the output (residual) for logarithm method will be orthogonal with the first derivative of the norm. If we use L2 norm, this tells us the output will be white spectrum. But for linear method, we can not get this property.

That is why we need logarithm method, or why we need a exponential function in Fourier domain. The exponential function can help us get a gradient of residual as a shift residual rather than a shift data.