Mixtures, envelopes and hierarchical duality

Check loss function

$$2\rho_\tau(x) = \begin{cases} 2\tau x & x \geq 0 \\ 2(\tau - 1)x & x < 0 \end{cases} = |x| + (2\tau - 1)x$$

Quantile Regression with Trend Filtering Problem

$$\underset{\theta \in \mathcal{R}^d}{\text{minimize}} \sum_{i=1}^{n} \{|y_i - \theta_i| + (2\tau - 1)(y_i - \theta_i)\} + \lambda ||D^{k+1}\theta||_1$$

Define

$$\begin{aligned} r_i &= y_i - \theta_i \\ f(r) &= |r| + (2\tau - 1)(r) \\ g(r) &= |r| = |r| + (2\tau - 1)(r) + -(2\tau - 1)(r) = f(r) + \kappa r = \gamma(r^2/2) \end{aligned}$$

So $\gamma(x) = g(\sqrt{2x}) = \sqrt{2x}$ is concave on $\mathcal{R}^+$.

Let $\gamma^*(u)$ be the concave dual of $\gamma(x)$, so that

$$\gamma(x) = \inf_u \{x^T u - \gamma^*(u)\}$$
$$\gamma^*(u) = \inf_x \{x^T u - \gamma(x)\}$$

Then we can write

$$\begin{aligned} f(r) + \kappa r =& \gamma\left(\frac{r^2}{2}\right) \\ f(r) + \kappa r =& \inf_u \left\{\frac{r^2 u}{2} - \gamma^*(u)\right\} \\ f(r) =& \inf_u \left\{\frac{r^2 u}{2} - \kappa r - \gamma^*(u)\right\} \\ f(r) =& \inf_u \left\{\frac{u}{2}\left(r - \frac{\kappa}{u}\right)^2 - \frac{\kappa^2}{2u} - \gamma^*(u)\right\} \\ f(r) =& \inf_u \left\{\frac{u}{2}\left(r - \frac{\kappa}{u}\right)^2 - \psi(u)\right\} \end{aligned}$$

where

$$\psi(u) = \frac{\kappa^2}{2u} + \gamma^*(u)$$

Then the likelihood $p(r) \propto e^{-f(x)}$ has an envelope representation as a variance-mean normal distribution with drift parameter κ.

$$p(r) \propto \exp(-f(x)) = \sup_u \left\{ \mathcal{N}(r|\kappa u^{-1}, u^{-1}) u^{-1/2} e^{\psi(u)} \right\}$$

and any optimal value of u as a function of r satisfies $\widehat{u}(r) \in \partial\gamma(r^2/2)$. So in the case that f is differentiable

$$\widehat{u}(r) = \frac{f'(r) + \kappa}{r}$$

We can re-write our minimization problem:

old formulation

$$\underset{\theta \in \mathcal{R}^d}{\text{minimize}} \sum_{i=1}^{n} \{|y_i - \theta_i| + (2\tau - 1)(y_i - \theta_i)\} + \lambda||D^{k+1}\theta||_1$$

new formulation

$$\underset{\theta \in \mathcal{R}^d}{\text{minimize}} \sum_{i=1}^{n} \inf_{u_i > 0} \left\{ \frac{u_i}{2} \left(y_i - \theta_i - \frac{1 - 2\tau}{u_i} \right)^2 - \psi(u_i) \right\} + \lambda||D^{k+1}\theta||_1$$

Algorithm Updates

$$u_i^{(t)} = \frac{sgn(y_i - \theta_i^{(t-1)})}{y_i - \theta_i^{(t-1)}}$$

$$z_i^{(t)} = y_i - (1 - 2\tau)/w_i^{(t)}$$

$$\theta^{(t)} = \underset{\theta}{\text{argmin}} \sum_{i=1}^{n} \frac{u_i^{(t)}}{2} (z_i - \theta_i)^2 + \lambda||D^{k+1}\theta||_1$$

1 FASTA

Fix $\nu = 1, .1, .001, .001$ want to minimize:

$$f(\eta, \theta) + g(\eta)$$

$$f(\eta, \theta) = \frac{1}{n} \tilde{\rho_\tau}(y_i - \theta_i) + \frac{1}{2\nu} ||\eta - D\theta||_2^2$$

$$g(\eta) = \lambda||\eta||_1$$

$$\begin{aligned} \rho_\tau(z) =& z[\tau - I(z < 0)] \\ =& \frac{1}{2}|z| + \frac{2\tau - 1}{2} z \end{aligned}$$

$$\rho'_\tau(z) = \begin{cases} \tau & z > 0 \\ \tau - 1 & z < 0 \end{cases}$$

$$\frac{\partial}{\partial\theta_i}\frac{1}{n}\rho_\tau(y_i-\theta_i) = \frac{-1}{n}\rho'_\tau(y_i-\theta_i) = \begin{cases} -\tau/n & y_i-\theta_i > 0 \\ (1-\tau)/n & y_i-\theta_i < 0 \end{cases}$$