Question

Saturday, November 30, 2019 1:12 PM

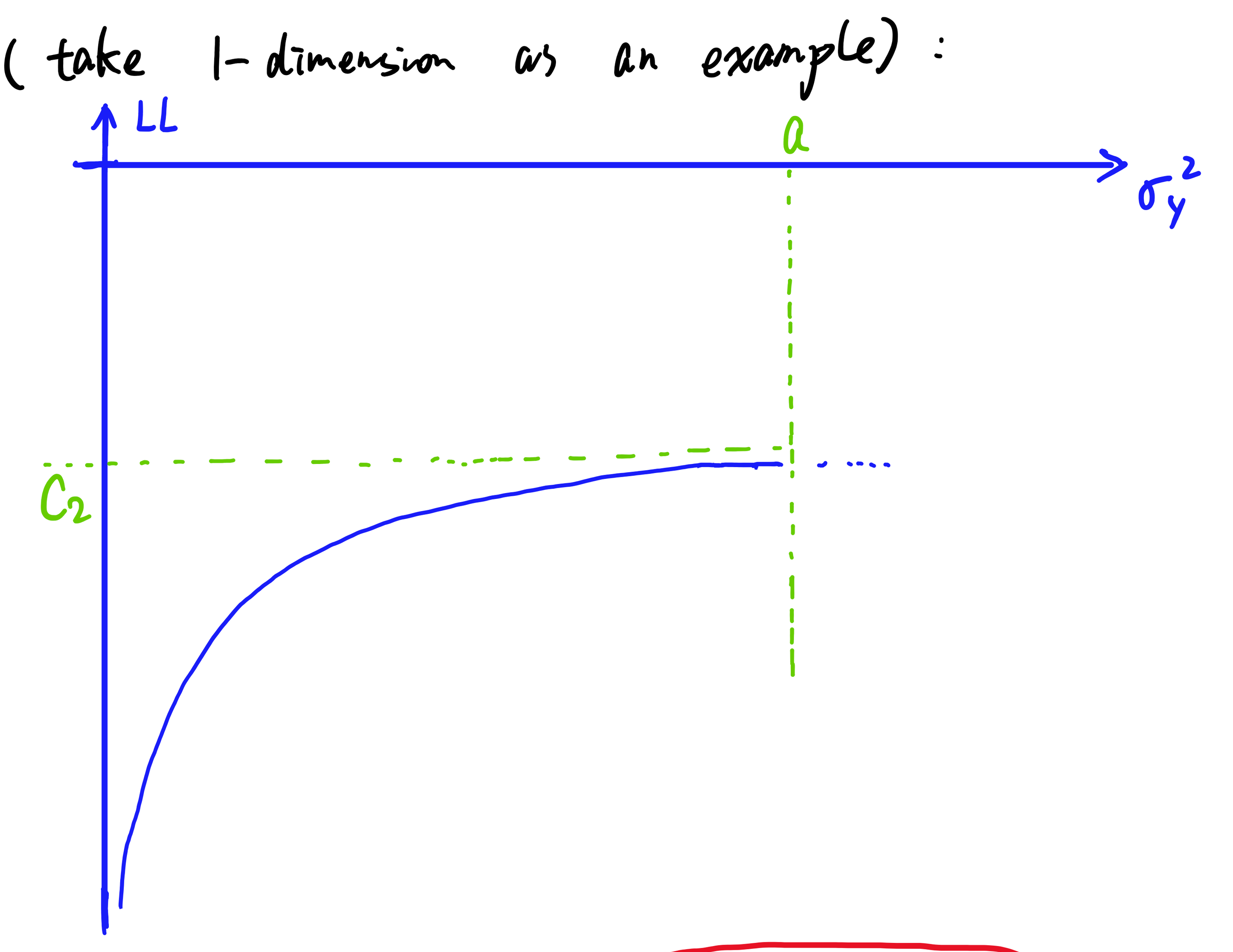
Hi Aaron. I met some problems when dealing with my research, and they are as following:

$$LL = -\frac{n-11}{2}\log\sigma_y^2 - \frac{n-11}{2}\log\sigma_T^2 - \frac{n-11}{2}\log\sigma_S^2 - \frac{\sum_{i=12}^{n}(y_i-T_i-S_i)^2}{2\sigma_y^2} - \frac{\sum_{i=12}^{n}(T_i-T_{i-1})^2}{2\sigma_T^2} - \frac{\sum_{i=12}^{n}\left(\sum_{j=0}^{11}S_{i-j}\right)^2}{2\sigma_S^2}$$

$$LOSS = \frac{\sum(I_{r_{X11}}-I_{r_{SSM}})^2}{\sigma_y^2} + \frac{\sum(T_{X11}-T_{SSM})^2}{\sigma_T^2} + \frac{\sum(S_{X11}-S_{SSM})^2}{\sigma_S^2}$$

$$LP = -\beta\cdot x^{\alpha}$$

And we already know the LL curve usually behaves like (take 1-dimension as an example):



I thought we can use $\boxed{f_{(x)} = -\frac{C_1}{x} - C_2}$ as an approximation $(0 < x \le a)$

Further, if $\alpha = 1$, for σ_y^2:

$$MAP_y = \underset{\sigma_y^2}{\arg\max}\,(LL_y + LP_y)$$

$$LPost_y = LL_y + LP_y = -\frac{C_{1y}}{\sigma_y^2} - C_{2y} - \beta_y\sigma_y^2$$

Let $\frac{\partial LPost_y}{\partial \sigma_y^2} = \frac{C_{1y}}{(\sigma_y^2)^2} - \beta_y = 0 \Rightarrow \hat{\sigma}_y^2 = \sqrt{\frac{C_{1y}}{\beta_y}}$

Suppose $(\sigma_y^2, \sigma_T^2, 1)^*$ is the result from grid search, usually speaking, $\sigma_y^{2*} \in [2, 10]$, then we have

$$2 \le \sqrt{\frac{C_{1y}}{\beta_y}} \le 10$$

$$\Rightarrow \frac{C_{1y}}{100} \le \beta_y \le \frac{C_{1y}}{4}$$

(I devided this interval into 10 parts and then use grid search to check which prior is better.)

Let $\beta_y = \frac{C_{1y}}{100} + k_y \cdot P_y$ where $P_y = \frac{1}{9}\left[\frac{C_{1y}}{4} - \frac{C_{1y}}{100}\right] = \frac{2}{75}C_{1y}$

$k_y = 0, 1, \dots 9.$

Similarly, $\beta_T = \frac{C_{1T}}{100} + k_T \cdot \frac{2}{75}C_{1T}$ (suppose $\sigma_T^2 \in [2, 10]$)

Therefore, if we can have reasonable C_{1y} and C_{1T} of each dataset, we can use grid search to find constants k_y & k_T which generate the lowest loss among 10 simulated datasets.

Now, we look at the expression $\boxed{LL_y = -\frac{C_{1y}}{\sigma_y^2} - C_{2y}}$

Back to our definition of LL, $C_{1y} = \frac{\sum_{i=12}^{n}(y_i-T_i-S_i)^2}{2}$

But C_{1y} will change given different σ_y^2, σ_T^2.

[I thought $\frac{\sum(y_i-T_i-S_i)^2}{2}$ & $\frac{\sum(T_i-T_{i-1})^2}{2}$ don't have too much difference before, cause visulizationally the decompositions' difference is not obvious. Thus, based on this assumption, I took the average among $\sigma_y^2 \in Int[1,10]$ & $\sigma_T^2 \in Int[1,10]$.

BUT, the result was not good.

Then I checked the values of $\frac{\sum(y_i-T_i-S_i)^2}{2}$ & $\frac{\sum(T_i-T_{i-1})^2}{2}$ from different variances' setting, and found the value is quite different actually although the decomposition looks similar.]

Therefore, my first question is: how to deal with C_{1y} and C_{1T}? (If my thoughts else is correct).

Question 2.

Assume my thoughts above are reasonable and we figure out the problem of C_{1y} & C_{1T}, then apparently we can do grid search of constants k_y & k_T over 10 simulated datasets, but I have one problem here.

If we choose the k_y & k_T which generate the lowest total loss, the result will be dominated by some datasets with greater magnitudes, longer lengths or something else. Therefore, we need to 'normalize' the loss for each dataset and then compute the total. But I have no Idea of how to normalize the loss for now.