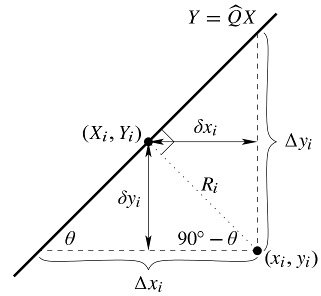
Definitions arising from new geometry

- Recall the figure to the right, which will define the geometry of our approximate WTLS solution
- Define Δx_i be x-distance between data point i and line, and Δy_i be y-distance between data point i and line
- Slope of line is $\widehat{Q} = \Delta y_i / \Delta x_i$ for all i
- Angle of line is $\theta = \tan^{-1} \widehat{Q}$
- Shortest distance between line and any given data point:

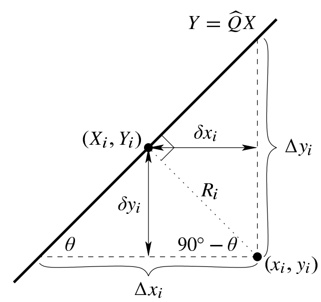
$$\begin{aligned} R_i &= \Delta y_i \cos\theta \\ &= \Delta y_i / \sqrt{1 + \tan^2\theta} = \Delta y_i / \sqrt{1 + \widehat{Q}^2} \end{aligned}$$



First form of cost function

- Let $\delta x_i = R_i \sin\theta$ and $\delta y_i = R_i \cos\theta$
- These are the x- and y-components of the perpendicular distance between data point i and the fitting line
- We then weigh our fitting cost function according to these variances
- Therefore, we define the approximate weighted total least squares (AWTLS) cost function as

$$\chi^2_{\text{AWTLS}} = \sum_{i=1}^{N} \frac{\delta x_i^2}{\sigma_{x_i}^2} + \frac{\delta y_i^2}{\sigma_{y_i}^2}$$



More useful form of cost function

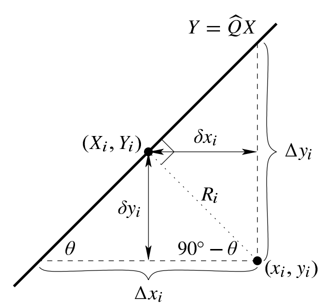
- Note that $\sin^2\theta = 1 - \cos^2\theta = \widehat{Q}^2/(1 + \widehat{Q}^2)$:

$$\delta x_i^2 = \left(\frac{\Delta y_i^2}{1+\widehat{Q}^2}\right)\left(\frac{\widehat{Q}^2}{1+\widehat{Q}^2}\right); \quad \delta y_i^2 = \left(\frac{\Delta y_i^2}{1+\widehat{Q}^2}\right)\left(\frac{1}{1+\widehat{Q}^2}\right)$$

- Since $\Delta y_i = y_i - \widehat{Q}x_i$

$$\chi^2_{\text{AWTLS}} = \sum_{i=1}^{N} \frac{(y_i - \widehat{Q}x_i)^2}{(1 + \widehat{Q}^2)^2} \left(\frac{\widehat{Q}^2}{\sigma_{x_i}^2} + \frac{1}{\sigma_{y_i}^2}\right)$$

- To verify that AWTLS approximates WTLS in some cases, note both cost functions equal when $\sigma_{x_i} = \sigma_{y_i}$
- However, they are not equal when $\sigma_{x_i} = k\sigma_{y_i}$, but this will be corrected in a later lesson



Summary

- Starting with geometry presented in last lesson, defined some quantities describing line and distances from data point to line in different directions
- Proposed new cost function in terms of these new definitions
- Then, rewrote cost function in terms of quantities available to us
- We are now ready to derive the solution that optimizes this cost function