# Inclusion of the angle in the Normalized Least Squares Objective Function

Zicheng Gao (zxg109)

October 10, 2016

We first define the $\cos(\theta)$ between two vectors $\delta, \gamma \in \mathbb{R}$

$$\cos(\theta) = \frac{\delta \cdot \gamma}{\parallel \delta \parallel \parallel \gamma \parallel}$$

Then we begin with the Normalized Least Squares formula for the two vectors.

$$\begin{aligned}
\text{NLS} &= \sum_i \left( \frac{\delta_i}{\parallel \delta \parallel} - \frac{\gamma_i}{\parallel \gamma \parallel} \right)^2 && (1) \\
&= \sum_i \left( \left( \frac{\delta_i}{\parallel \delta \parallel} \right)^2 - 2\frac{\delta_i \gamma_i}{\parallel \delta \parallel \parallel \gamma \parallel} + \left( \frac{\gamma_i}{\parallel \gamma \parallel} \right)^2 \right) && (2) \\
&= \sum_i \left( \frac{\delta_i}{\parallel \delta \parallel} \right)^2 - 2\sum_i \frac{\delta_i \gamma_i}{\parallel \delta \parallel \parallel \gamma \parallel} + \sum_i \left( \frac{\gamma_i}{\parallel \gamma \parallel} \right)^2 && (3)
\end{aligned}$$

As $\parallel \alpha \parallel^2 = \alpha \cdot \alpha = \sum_i \alpha_i^2$ and $\sum_i \alpha_i \beta_i = \alpha \cdot \beta$ and $\parallel \alpha \parallel$ is a scalar, we can rewrite this:

$$\begin{aligned}
&= \sum_i \left( \frac{\delta_i^2}{\parallel \delta \parallel^2} \right) - 2\sum_i \frac{\delta_i \gamma_i}{\parallel \delta \parallel \parallel \gamma \parallel} + \sum_i \left( \frac{\gamma_i^2}{\parallel \gamma \parallel^2} \right) && (4) \\
&= \frac{\sum_i \delta_i^2}{\parallel \delta \parallel^2} - 2\frac{\sum_i \delta_i \gamma_i}{\parallel \delta \parallel \parallel \gamma \parallel} + \frac{\sum_i \gamma_i^2}{\parallel \gamma \parallel^2} && (5) \\
&= \frac{\delta \cdot \delta}{\delta \cdot \delta} - 2\frac{\delta \cdot \gamma}{\parallel \delta \parallel \parallel \gamma \parallel} + \frac{\gamma \cdot \gamma}{\gamma \cdot \gamma} && (6) \\
&= 1 - 2\cos(\theta) + 1 && (7) \\
&= 2(1 - \cos(\theta)) && (8)
\end{aligned}$$

Thus it is clear that the Normalized Least Squares method in fact utilizes the angle difference between the vectors. Compare to the quadratically wrapped shape objective function:

$$(1 - \cos(\theta))^2$$

And to the sharpness-modified shape objective function:

$$(1 - \cos(\theta))^2 + 1 - e^{-Q(1-\cos(x))^2}$$

...where Q determines the "sharpness" of the slope near the minimum - not that this would truly impact the performance of the regression.

As all these "wrapping functions" around the angle preserve monoticity, there is no functional difference between their usage beyond computational requirements if only the value is used (that is, the gradient is not used).

Overall, we can see that the NLS and shape objective function are deeply related.