Gauss-Newton

Joshua Clark, Kyle Grinstead, Matthew Le

November 23, 2014

1 Proof

The reason that $\beta = \beta - R^{-1}Q^T r$ is equivalent to $\beta = \beta - (J^T J)^{-1} J^T r$ is as follows:

$$J = QR$$

$$\beta = \beta - ((QR)^T (QR))^{-1} (QR)^T r$$

$$Q^T Q = I$$

$$\beta = \beta - (R^T I R)^{-1} R^T Q^T r$$

$$\beta = \beta - R^{-1} R^{T^{-1}} R^T Q^T r$$

$$R^{T^{-1}} R^T = I$$

$$\beta = \beta - R^{-1} I Q^T r$$

$$\beta = \beta - R^{-1} Q^T r$$

2 Algorithm Benefits

Why is this new algorithm ($\beta = \beta - R^{-1}Q^T r$) better than the original? Well, in short, because it reduces the error by minimizing the condition number. When using the original equation, every time J or J^T is used, it multiplies the condition number by J's condition number. J is used three times in that equation so the error will be extremely magnified. When using Q and R, however, there is no magnification of error. The condition number for Q (and Q^T) is 1 due to the fact that it is an orthogonal matrix. The condition number for R (and R^{-1}) is the same as the condition number for J. The end result is that the new algorithm will have the minimum amount of error allowed by the original matrix's condition number.