Why is it justified to modify the algorithm to set **β** to **β** – **R-1QT** ?

We go from β = β – (JTJ)-1JTr to β = β – R-1QTr.

Therefore we can say:

(JTJ)-1JT = R-1QT

We are performing QR on the Jacobi matrix J, so we know J = QR.

Therefore:

(RTQTQR)-1RTQT = R-1QT

Since Q is an orthogonal matrix, we can say that QTQ = I.

(RTIR)-1RTQT = R-1QT

We can right multiply both sides by Q:

(RTR)-1RTQTQ = R-1QTQ

(RTR)-1RTI = R-1I

(RTR)-1RT = R-1

What is the benefit of modifying the algorithm in this way?

This algorithm is more efficient. When you left multiply numerous times by matrices with unbounded condition numbers (error amplification), the algorithm is not as stable. Doing the algorithm using QR is much more efficient because, since Q is an orthogonal matrix we know that the condition number k(Q) = 1. This way, the condition number/ error amplification is solely depended on one matrix R 🡪 k(R)=k(A) (A being J in this case). Because of this, the modified algorithm has less error than the original algorithm.