OLS estimate of $\boldsymbol{\gamma}$ is

$$\hat{\boldsymbol{\gamma}} = \left(\boldsymbol{Q}^\top\boldsymbol{Q}\right)^{-1}\boldsymbol{Q}^\top\boldsymbol{y} = \boldsymbol{Q}^\top\boldsymbol{y},$$

which is trivial to compute. It is equally easy to compute the fitted values $\boldsymbol{Q}\hat{\boldsymbol{\gamma}}$ and the residuals

$$\hat{\boldsymbol{u}} = \boldsymbol{y} - \boldsymbol{Q}\hat{\boldsymbol{\gamma}} = \boldsymbol{y} - \boldsymbol{Q}\boldsymbol{Q}^\top\boldsymbol{y}. \tag{1.36}$$

Thus, if we are simply interested in residuals and/or fitted values, we do not need to compute $\hat{\boldsymbol{\beta}}$ at all.

Notice from (1.36) that the projection matrices $\boldsymbol{P}_X$ and $\boldsymbol{M}_X$ are equal to $\boldsymbol{Q}\boldsymbol{Q}^\top$ and $\mathbf{I} - \boldsymbol{Q}\boldsymbol{Q}^\top$, respectively. The simplicity of these expressions follows from the fact that $\boldsymbol{Q}$ forms an orthonormal basis for $\mathcal{S}(\boldsymbol{X})$. Geometrically, nothing would change in any of the figures we have drawn if we used $\boldsymbol{Q}$ instead of $\boldsymbol{X}$ as the matrix of regressors, since $\mathcal{S}(\boldsymbol{Q}) = \mathcal{S}(\boldsymbol{X})$. If we were to show the columns of $\boldsymbol{Q}$ in the figures, each column would be a point in $\mathcal{S}(\boldsymbol{X})$ located on the unit sphere (i.e., the sphere with radius one centered at the origin) and at right angles to the points representing the other columns of $\boldsymbol{Q}$.

In order to calculate $\hat{\boldsymbol{\beta}}$ and $(\boldsymbol{X}^\top\boldsymbol{X})^{-1}$, which, along with the residuals and the fitted values, allow us to calculate all the main quantities of interest, we make use of the facts that $\hat{\boldsymbol{\beta}} = \boldsymbol{R}^{-1}\hat{\boldsymbol{\gamma}}$ and

$$\left(\boldsymbol{X}^\top\boldsymbol{X}\right)^{-1} = \left(\boldsymbol{R}^\top\boldsymbol{Q}^\top\boldsymbol{Q}\boldsymbol{R}\right)^{-1} = \left(\boldsymbol{R}^\top\boldsymbol{R}\right)^{-1} = \boldsymbol{R}^{-1}(\boldsymbol{R}^{-1})^\top.$$

Thus, once we have computed $\boldsymbol{R}^{-1}$, we can very easily calculate the least squares estimates $\hat{\boldsymbol{\beta}}$ and their estimated covariance matrix (see Chapter 2). Since $\boldsymbol{R}$ is a triangular matrix, its inverse is very easily and cheaply computed; we do not even have to check for possible singularity, since $\boldsymbol{R}$ will fail to have full rank only if $\boldsymbol{X}$ does not have full rank, and that will already have shown up and been dealt with when we formed $\boldsymbol{Q}$ and $\boldsymbol{R}$.

The most costly part of these procedures is forming the matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$ from $\boldsymbol{X}$. This requires a number of arithmetic operations that is roughly proportional to nk^2. Forming the matrix of sums and cross-products, which is the first step for methods based on solving the normal equations, also requires a number of operations proportional to nk^2, although the factor of proportionality is smaller. Thus linear regression by any method can become expensive when the number of regressors is large and/or the sample size is very large. If one is going to calculate many regressions using the same large data set, it makes sense to economize by doing the expensive calculations only once. Many regression packages allow users first to form the matrix of sums of squares and cross-products for all the variables in a data set and then to calculate estimates for a variety of regressions by retrieving the relevant rows and columns and using normal equation methods. If this approach is used, it is particularly important that the data be scaled so that the various regressors are not too dissimilar in mean and variance.