QR Decomposition

- The columns of Q consists of an orthonormal basis for the column space of X.
- Q is orthogonal matrix of $n \times d$, satisfying $Q^T Q = I$.
- R is upper triangular matrix of $d \times d$, full-ranked.
- $X^T X = (QR)^T (QR) = R^T Q^T QR = R^T R$

The least square solutions are

$$
\begin{aligned}
\widehat{\beta} &= (X^T X)^{-1} X^T \mathbf{y} \\
&= R^{-1} R^{-T} R^T Q^T \mathbf{y} = R^{-1} Q^T \mathbf{y} \\
\hat{\mathbf{y}} &= X \hat{\beta} \\
&= (QR)(R^{-1} Q^T \mathbf{y}) \\
&= QQ^T \mathbf{y}.
\end{aligned}
$$

QR Algorithm for Normal Equations

Regard $\widehat{\beta}$ as the solution for linear equations system:

$$R\boldsymbol{\beta} = Q^T\mathbf{y}.$$

1. Conduct QR decomposition of $X = QR$. (Gram-Schmidt Orthogonalization)
2. Compute $Q^T\mathbf{y}$.
3. Solve the triangular system $R\boldsymbol{\beta} = Q^T\mathbf{y}$.

The computational complexity: nd^2

Cholesky Decomposition Algorithm

For any positive definite square matrix A, we have

$$A = RR^T,$$

where R is a lower triangular matrix of full rank.

1. Compute X^TX and $X^T\mathbf{y}$.
2. Factoring $X^TX = RR^T$, then $\hat{\beta} = (R^T)^{-1}R^{-1}X^T\mathbf{y}$
3. Solve the triangular system $R\mathbf{w} = X^T\mathbf{y}$ for $\mathbf{w}$.
4. Solve the triangular system $R^T\boldsymbol{\beta} = \mathbf{w}$ for $\boldsymbol{\beta}$.

The computational complexity: $d^3 + nd^2/2$ (can be faster than QR for small d, but can be less stable)

$$\text{Var}(\hat{\mathbf{y}}_0) = \text{Var}(\mathbf{x}_0^T\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{x}_0^T(R^T)^{-1}R^{-1}\mathbf{x}_0).$$