Gram–Schmidt Orthogonalization

Given independent vectors $a_1, a_n, \ldots, a_n$ the Gram–Schmidt orthogonalization algorithm constructs orthonomal vectors $q_1, q_2, \ldots, q_n$ as follows

$$\begin{aligned}
\tilde{q}_1 &= a_1 & q_1 &= \tilde{q}_1/\|\tilde{q}_1\| \\
\tilde{q}_2 &= a_2 - q_1(q_1 \cdot a_2) & q_2 &= \tilde{q}_2/\|\tilde{q}_2\| \\
\tilde{q}_3 &= a_3 - q_1(q_1 \cdot a_3) - q_2(q_2 \cdot a_3) & q_3 &= \tilde{q}_3/\|\tilde{q}_3\| \\
&\vdots & &\vdots \\
\tilde{q}_n &= a_n - q_1(q_1 \cdot a_n) - q_2(q_2 \cdot a_n) - \cdots - q_{n-1}(q_{n-1} \cdot a_n) & q_n &= \tilde{q}_n/\|\tilde{q}_n\|.
\end{aligned}$$

Note that

$$\operatorname{span}\{a_1, \ldots, a_k\} = \operatorname{span}\{q_1, \ldots, q_k\} \qquad \text{for every } k = 1, 2, \ldots, n.$$

Let

$$A = \left[a_1 \middle| a_2 \middle| \cdots \middle| a_n \right] \qquad \text{and} \qquad Q = \left[q_1 \middle| q_2 \middle| \cdots \middle| q_n \right].$$

Since $q_i \in \operatorname{span}\{q_1, \ldots, q_k\}^\perp = \operatorname{span}\{a_1, \ldots, a_k\}^\perp$ for every $i > k$, then

$$q_i \cdot a_j = 0 \qquad \text{for every } i > j.$$

It follows that

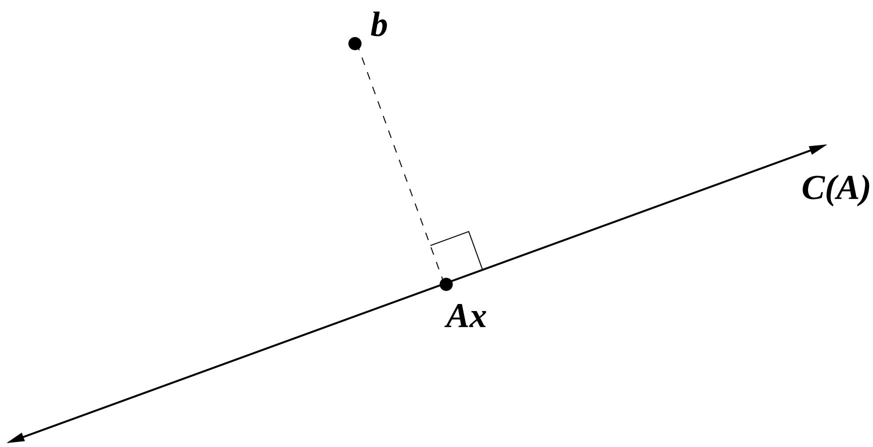
$$R = Q^T A = \begin{bmatrix} q_1^T \\ \hline q_2^T \\ \hline \vdots \\ \hline q_n^T \end{bmatrix} \left[a_1 \middle| a_2 \middle| \cdots \middle| a_n \right] = \begin{bmatrix} q_1 \cdot a_1 & q_1 \cdot a_2 & \cdots & q_1 \cdot a_n \\ 0 & q_2 \cdot a_2 & \cdots & q_2 \cdot a_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_n \cdot a_n \end{bmatrix}$$

is upper triangular. Moreover,

$$\begin{aligned}
QR &= \left[q_1 \middle| q_2 \middle| \cdots \middle| q_n \right] \begin{bmatrix} q_1 \cdot a_1 & q_1 \cdot a_2 & \cdots & q_1 \cdot a_n \\ 0 & q_2 \cdot a_2 & \cdots & q_2 \cdot a_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_n \cdot a_n \end{bmatrix} \\
&= \left[q_1(q_1 \cdot a_1) \middle| q_1(q_1 \cdot a_2) + q_2(q_2 \cdot a_2) \middle| \cdots \middle| q_1(q_1 \cdot a_n) + q_2(q_2 \cdot a_n) + \cdots + q_n(q_n \cdot a_n) \right] \\
&= \left[a_1 \middle| a_2 \middle| \cdots \middle| a_n \right] = A.
\end{aligned}$$

Least Squares Using $A = QR$

The least squares approximation is the point x such that $\|Ax - b\|$ is minimized. This happens when $Ax - b$ is orthogonal to $\mathcal{C}(A)$.



Since $\mathcal{C}(A) = \mathcal{C}(Q)$, then we are looking for x such that

$$q_i \cdot (Ax - b) = 0 \qquad \text{for } i = 1, 2, \ldots, n.$$

Therefore

$$\begin{aligned} Q^T(Ax - b) &= \mathbf{0} \\ Q^T Ax &= Q^T b \\ Q^T(QR)x &= Q^T b \\ (Q^T Q)Rx &= Q^T b \\ Rx &= Q^T b. \end{aligned}$$

Since R is upper triangular, this system may be easily solved by back substitution.