Introduction Our first projection formula started with an orthonormal set $\{q_1, q_2, \ldots, q_m\} \subseteq \mathbb{R}^n$. We defined $Y := \text{span}\{q_1, q_2, \ldots, q_m\}$ and defined, for each $x \in \mathbb{R}^n$,

$$\text{proj}_Y x := (x \bullet q_1)q_1 + \cdots + (x \bullet q_m)q_m = \sum_{j=1}^{m} q_m q_m^T x = QQ^T x,$$

where Q is the $n \times m$ matrix with columns $q_1, q_2, \ldots, q_m$: $Q = (\, q_1 \quad q_2 \quad \cdots \quad q_m \,)$. We observed that $\text{proj}_Y x$ is in Y and we have abused notation by writing $\text{proj}_Q x$ instead of $\text{proj}_Y x$, because it is easier to write and follow than the "correct" $\text{proj}_{\text{span}\, Q}\, x$. We also showed that $x - \text{proj}_Y x \perp Y$ which means $(x - \text{proj}_Y x) \bullet y = 0$ for all $y \in Y$. We also wrote this as $x - \text{proj}_Y x \in Y^\perp$. We noticed that $Q^T Q = I_m$, while QQ^T is the projection-matrix and its size is $n \times n$.

This note will give a list of similar formulas for projection matrices when we do not begin with an orthonormal set of vectors.

Starting with a basis Suppose we are given (or find) a basis $\{v_1, v_2, \ldots, v_m\}$ of a subspace $Y \subseteq \mathbb{R}^n$. We saw how to begin with the $n \times m$ matrix $V = (\, v_1 \quad v_2 \quad \cdots \quad v_m \,)$ whose columns v_j, in $\mathbb{R}^n$, are linearly independent (so $m \le n$) and construct a matrix $Q = (\, q_1 \quad q_2 \quad \cdots \quad q_m \,)$ whose columns form an orthonormal set having the same span as the columns of V: we used the Gram–Schmidt–2573H method to get a matrix $W = (\, w_1 \quad w_2 \quad \cdots \quad w_m \,)$ whose columns form an orthogonal, not necessarily orthonormal, set having the same span as the columns of V. We then solved for the v_k in terms of the w_k, and noticed that the formulas

$$v_k = \sum_{j=1}^{k} \frac{v_k \bullet w_j}{|w_j|^2}\, w_j,\ k = 1, \ldots, m$$

could be rewritten

$$v_k = \sum_{j=1}^{k} (v_k \bullet q_j)\, q_j,\ k = 1, \ldots, m, \quad \text{where each } \ q_j := w_j/|w_j|.$$

and this finally gave us the QR decomposition: $V = QR$, where $R = (r_{ij})$ is upper triangular and invertible, the columns of Q form an orthonormal set and $\text{Im}\ Q = \text{Im}\ V$ $(r_{ij} = v_j \bullet q_i$ if $i \le j$ and $r_{ij} = 0$ if $i > j)$.

This gives us the same formula as our first one, if we write $Y = \text{Im}\ V$ because $Y = \text{Im}\ Q$ is also true. We can then write $\text{proj}_Y x = QQ^T x$ for each $x \in \mathbb{R}^n$. **Remark:** To form Q we have to divide by some "square–root" quantities. If all we want is QQ^T we can avoid square roots. We recall that $QQ^T = \sum_{j=1}^{m} q_j q_j^T$, so

$$(*) \qquad QQ^T = \sum_{j=1}^{m} q_j q_j^T = \sum_{j=1}^{m} \frac{w_j w_j^T}{|w_j|^2}.$$

This formula shows that the projection–matrix has rational–number entries if the original matrix did.

We would like to write the projection in terms of V instead of Q. We will use the QR decomposition. Thus $V = QR$ and so $V^T = R^T Q^T$. But then we notice that $V^T V = R^T Q^T Q R = R^T R$ because $Q^T Q = I_m$. Since R is invertible, so is R^T (we have $R^{-1}R = RR^{-1} = I_m$ so $(R^{-1})^T R^T = (RR^{-1})^T = I_m^T = I_m$ and thus $(R^T)^{-1} = (R^{-1})^T$). The product of invertible matrices is invertible (we have $(AB)^{-1} = B^{-1}A^{-1}$) so $V^T V (= R^T R)$ is invertible. This does not mean that V is invertible!

We have $V = QR$, so $Q = VR^{-1}$ and $Q^T = (R^T)^{-1} V^T$. Hence

$$QQ^T = VR^{-1}(R^T)^{-1}V^T = V(R^T R)^{-1}V^T = V(V^T V)^{-1}V^T.$$

This gives the formula we want:

$$\text{proj}_Y x = \text{proj}_V x = QQ^T x = V(V^T V)^{-1} V^T x \ \ \text{(this is Bretscher's Fact 5.4.8).}$$

Starting with a matrix Suppose we are given an $n \times p$ matrix A with and we want an $n \times n$ matrix P such that $Px = \text{proj}_{\text{Im } A} x$ for all $x \in \mathbb{R}^n$. We recall that $\text{Im } A$ is a subspace of $\mathbb{R}^n$, so its dimension is at most n. We also know that the dimension of $\text{Im } A$ is the rank of A, which is the number of linearly independent columns of A and hence is also at most p. We will let m denote the rank of A, so $m \le \min\{n, p\}$.

There are two approaches we can take. The first is easiest to follow. The second may be faster.

first approach We carry out Gauss–Jordan elimination on A: $(\ A \mid I\ ) \to (\ R \mid M\ )$, and extract from A the columns $v_1 := Ae_{j_1}, \ldots, v_m := Ae_{j_m}$ that correspond to the columns of R with leading ones. We then define $V := (v_1 \cdots v_m)$ and follow the "Starting with a basis" steps. We have two choices: (1) use Gram–Schmidt–2573H to find W and then QQ^T, or (2) we can compute $V(V^TV)^{-1}V^T$.

second approach We can apply the Gram–Schmidt–2573H process to the columns of A directly, even if we do not know that the columns are linearly independent. There is an important difference, though! If some column of A, say column k, is a linear combination of preceding columns, then w_k will be zero! We just ignore that column and keep going. We will find exactly $m = \operatorname{rank} A$ orthogonal columns (we have to go thru all the columns of A tho [unless we accumulate $\min\{n, p\}$ w_k's] because we do not know m in advance). We use the non-zero columns w_j, and compute QQ^T as in $(*)$.

Review of the Gram–Schmidt–2573H method Given vectors $\{v_1, v_2, \ldots, v_m\}$ that form a linearly independent set we constructed vectors $\{w_1, w_2, \ldots, w_m\}$ as follows:

$$
(\dagger) \qquad
\begin{aligned}
w_1 :=& c_1 v_1, \quad \text{where } c_1 \ne 0; \\
w_k :=& c_k \left(v_k - \sum_{j=1}^{k-1} \frac{v_k \bullet w_j}{|w_j|^2}\, w_j \right), \quad \text{where } 1 < k \le m \text{ and } c_k \ne 0.
\end{aligned}
$$

The c_k will usually be used when $V = (v_1\ v_2\ \cdots\ v_m)$ has integer entries. Then c_k will be chosen to be the least common multiple of the denominators $|w_j|^2$ which results in vectors w_k with integer entries. The same approach can be used when V has rational entries.

As a matter of form, let's calculate $w_i \bullet w_k$ (when $i < k$). We assume that when $j < k$ and $i < k$ and $j \ne i$ then $w_j \bullet w_i = 0$. This is certainly true when $k = 2$! At any rate,

$$
\begin{aligned}
w_k \bullet w_i =& c_k \left(v_k \bullet w_i - \sum_{j=1}^{k-1} \frac{v_k \bullet w_j}{|w_j|^2}\, w_j \bullet w_i \right) \\
=& c_k \left(v_k \bullet w_i - \frac{v_k \bullet w_i}{|w_i|^2}\, w_i \bullet w_i \right) = 0.
\end{aligned}
$$

The summation disappears because $w_j \bullet w_i = 0$ except when $j = i$. Thus "true for $k = 2$" becomes "true for $k = 3$" and so on, as long as the k's last. We have shown that the w_k are mutually orthogonal.

When we carry out the work of finding the w_k it is a good idea to keep a list of the values of $|w_k|^2$ along with the w_k. The c_k do not have to be kept, except for theoretical purposes.

After finding the matrix $W = (w_1\ w_2\ \cdots\ w_m)$ we can form an orthonormal matrix by dividing each column of W by its length: $Q := (w_1/|w_1|\ w_2/|w_2|\ \cdots\ w_m/|w_m|) =: WD = (q_1\ q_2\ \cdots\ q_m)$. Then $Q^TQ = I_m$. Here D is a "diagonal" matrix, having entries $D_{ii} = 1/|w_i| > 0$ and all other entries zero, so D is invertible.

The QR decomposition The w_k were defined in the equations $(\dagger)$. We can solve those equations for the v_k:

$$
(\dagger\dagger) \qquad
\begin{aligned}
v_1 =& \frac{1}{c_1} w_1, \\
v_k =& \frac{1}{c_k} w_k + \sum_{j=1}^{k-1} \frac{v_k \bullet w_j}{|w_j|^2}\, w_j, \quad \text{where } 1 < k \le m.
\end{aligned}
$$

If we put $k=1$ in the second formula, we get a sum from 1 to zero, which is an empty sum. If we regard empty sums as zero, which is a standard convention, we can (and shall) use the second formula for $1 \le k \le m$.

Then

$$v_k \bullet w_k = \left(\frac{1}{c_k} w_k + \sum_{j=1}^{k-1} \frac{v_k \bullet w_j}{|w_j|^2}\, w_j \right) \bullet w_k = \frac{1}{c_k} w_k \bullet w_k + \sum_{j=1}^{k-1} \frac{v_k \bullet w_j}{|w_j|^2}\, w_j \bullet w_k = \frac{1}{c_k} w_k \bullet w_k = \frac{1}{c_k} |w_k|^2$$

since $w_j \bullet w_k = 0$ if $j < k$. Hence $\dfrac{1}{c_k} w_k = \dfrac{v_k \bullet w_k}{|w_k|^2}\, w_k$ and therefore we can redo the formulas (††) because the term $\dfrac{1}{c_k} w_k = \dfrac{v_k \bullet w_k}{|w_k|^2}\, w_k$ fits the format of the sum in (††) for $j = k$:

$$v_k = \sum_{j=1}^{k} \frac{v_k \bullet w_j}{|w_j|^2}\, w_j, \quad \text{for } 1 \le k \le m.$$

Let us notice that $\dfrac{v_k \bullet w_j}{|w_j|^2}\, w_j = \left(v_k \bullet \dfrac{w_j}{|w_j|} \right) \dfrac{w_j}{|w_j|}$ and recall that $\frac{w_j}{|w_j|} = q_j$, $1 \le j \le m$. Then we can re–redo the formulas (††):

$$(\dagger\dagger\dagger) \qquad \begin{aligned} v_k &= \sum_{j=1}^{k} \left(v_k \bullet \frac{w_j}{|w_j|} \right) \frac{w_j}{|w_j|} \\ &= \sum_{j=1}^{k} (v_k \bullet q_j)\, q_j, \quad \text{for } 1 \le k \le m. \end{aligned}$$

These equations tell us that each v_k is a linear combination of columns q_j, $1 \le j \le k$, of the $n \times m$ matrix Q. If we let the column vector $r_k = (r_{1k}, r_{2k}, \ldots, r_{mk})$, where $r_{ik} = \begin{cases} v_k \bullet q_i & \text{for } 1 \le i \le k \\ 0 & \text{for } i > k \end{cases}$, then $v_k = Qr_k$ for $1 \le i \le k$. Finally, when we let R be the $m \times m$ matrix with columns r_k, we see that

$$V = QR.$$

We note that $r_{kk} = q_k \bullet v_k = \frac{|w_k|}{c_k} \ne 0$, so R is invertible (since R is upper triangular).

We could have taken an alternate approach. Since $v_k = \sum_{j=1}^{k} (v_k \bullet q_j)\, q_j$, $q_i \bullet v_k = 0$ if $i > k$. Therefore $R = (q_i \bullet v_j) = (q_i^T v_j) = Q^T V$ (column i of Q, transposed, is row i of Q^T).

Examples Here are three matrices V: $\begin{pmatrix} -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 & 1 \\ -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$ and $\begin{pmatrix} 1 & -3 & -4 \\ -2 & 3 & 6 \\ 1 & -1 & -4 \\ 1 & 1 & 1 \end{pmatrix}$. Here are two matrices W: $\begin{pmatrix} 1 & -12 & 51 \\ -2 & 3 & 2 \\ 1 & 2 & -97 \\ 1 & 16 & 50 \end{pmatrix}$ and $\begin{pmatrix} 1 & 3 & 3 \\ -2 & 1 & 1 \\ 1 & -4 & 1 \\ 1 & 3 & -2 \end{pmatrix}$. Match them with their V's.

Here are three matrices QQ^T: $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2/3 & -1/3 & -1/3 \\ 0 & -1/3 & 2/3 & -1/3 \\ 0 & -1/3 & -1/3 & 2/3 \end{pmatrix}$, $\begin{pmatrix} 55/82 & -15/41 & -21/82 & -6/41 \\ -15/41 & 73/123 & -35/123 & -20/123 \\ -21/82 & -35/123 & 197/246 & -14/123 \\ -6/41 & -20/123 & -14/123 & 115/123 \end{pmatrix}$ and $\begin{pmatrix} 3/4 & -1/4 & -1/4 & -1/4 \\ -1/4 & 3/4 & -1/4 & -1/4 \\ -1/4 & -1/4 & 3/4 & -1/4 \\ -1/4 & -1/4 & -1/4 & 3/4 \end{pmatrix}$.

Match them with their V's. Find the matrices R such that $V = QR$.
It is easier to find $D^{-1}R = D^{-1}Q^TV = D^{-1}DW^TV = W^TV$
(see page 2, fourth and fifth lines from the bottom).